



**EVALUACIÓN DEL EFECTO DE LAS VARIABLES:  
GENERACIÓN, APORTES Y DEMANDA DE ENERGÍA DEL  
SISTEMA INTERCONECTADO NACIONAL EN EL PRECIO  
DE BOLSA NACIONAL DE ENERGÍA EN COLOMBIA  
PERIODO (2000 – 2014)**

**SANDRA ISABEL SABOGAL FLORIDO**

**FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES**  
DEPARTAMENTO DE CIENCIAS BÁSICAS  
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA  
BOGOTÁ D.C.  
2016



**EVALUACIÓN DEL EFECTO DE LAS VARIABLES:  
GENERACIÓN, APORTES Y DEMANDA DE ENERGÍA DEL  
SISTEMA INTERCONECTADO NACIONAL EN EL PRECIO  
DE BOLSA NACIONAL DE ENERGÍA EN COLOMBIA  
PERIODO (2000 – 2014)**

**SANDRA ISABEL SABOGAL FLORIDO**

Trabajo de grado para optar al título de Especialista en Estadística Aplicada

Asesor:  
Wilmer Pineda Ríos  
Docente

**FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES**  
DEPARTAMENTO DE CIENCIAS BÁSICAS  
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA  
BOGOTÁ D.C.  
2016

Nota de Aceptación

---

---

---

---

---

---

---

Firma del presidente del jurado

---

Firma del Jurado

---

Firma del Jurado

Bogotá, D.C 29 Enero del 2016

Las Directivas de la Universidad de  
Los Libertadores, los jurados calificadores y el cuerpo  
Docente no son responsables por los  
criterios e ideas expuestas en el presente documento.  
Estos corresponden únicamente a los autores.

## **AGRADECIMIENTOS**

Agradezco primero a Dios por permitirme concluir satisfactoriamente este nuevo proyecto académico. Al Profesor Wilmer Pineda, mi asesor, docente en la Universidad Santo Tomás, por su apoyo en varias etapas de la especialización, por sus múltiples contribuciones en el desarrollo y culminación del presente trabajo.

Agradezco igualmente a Giovanni N. por renovar en mí los deseos de regresar a la academia y apoyar mi amor por los números, a Shirley M. que me brindó su apoyo en todo momento tanto en lo académico como en lo personal y a mi familia con la que siempre puedo contar.

## TABLA DE CONTENIDO

	pág.
RESUMEN	12
1. INTRODUCCIÓN	14
2. PLANTEAMIENTO DEL PROBLEMA O PREGUNTA DE INVESTIGACIÓN	15
3. OBJETIVO GENERAL	15
4. OBJETIVOS ESPECÍFICOS	15
5. JUSTIFICACIÓN	16
6. MARCO DE REFERENCIA	17
6.1 ESTRUCTURA DEL MERCADO	17
6.2 ORGANISMOS DE ADMINISTRACIÓN Y CONTROL DEL MERCADO ELÉCTRICO COLOMBIANO	18
6.3 MERCADO ELÉCTRICO COLOMBIANO	18
6.3.1 Operaciones comerciales	18
6.3.2 Despacho ideal	19
6.4 DEFINICIONES	19
6.4.1 Precio de bolsa nacional	19
6.4.1.1 Cálculo del precio de bolsa nacional	20
6.4.2 Demanda	20
6.4.2.1 Demanda de energía	20
6.4.2.2 Demanda total doméstica	20
6.4.3 Generación	20
6.4.3.1 Generación real	20
6.4.3.2 Generación neta	20
6.5 ANÁLISIS DE DATOS FUNCIONALES (FDA)	20
6.5.1 Suavizado de funciones mediante funciones B-splines	21
6.5.2 Análisis descriptivo en datos funcionales	24
6.5.3 Herramientas gráficas del análisis de datos funcionales	24
6.5.4 Modelo de regresión funcional	26

	pág.
7. METODOLOGÍA	28
7.1 TIPO DE ESTUDIO	28
7.2 MÉTODO	29
7.3 VARIABLES DE ESTUDIO Y UNIDADES DE ANÁLISIS	29
7.4 PROCEDIMIENTO	30
7.4.1 Fase 1. Obtención de la información	30
7.4.2 Fase 2. Depuración de datos	31
7.4.3 Fase 3. Análisis de la información	31
7.5 ASPECTOS ÉTICOS	31
8. RESULTADOS	32
8.1 DATOS DISCRETIZADOS	32
8.2 DATOS SUAVIZADOS Y MEDIA FUNCIONAL	34
8.3 DESVIACIÓN ESTÁNDAR	38
8.4 IDENTIFICACIÓN DE SERIES ATÍPICAS	42
8.5 REGRESIÓN FUNCIONAL CONCURRENTES	46
9. CONCLUSIONES	51
10. DISCUSIÓN	54
REFERENCIAS	56
ANEXOS	58

## LISTA DE FIGURAS

	pág.
Figura 1. Flujo de información para el despacho ideal – Determinación del precio de bolsa	19
Figura 2. Comportamiento trimestral del PIB y la demanda de energía	39
Figura 3. Composición de la generación del SIN en 2014 (GWh)	41
Figura 4. Evolución histórica de los aportes energéticos al SIN 2000-2011	46



## LISTA DE GRÁFICAS

	pág.
Gráfica 1. Diagrama de caja funcional	25
Gráfica 2. Diagrama HDR funcional	26
Gráfica 3. Datos discretizados del precio de bolsa 2000 - 2014	32
Gráfica 4. Datos discretizados de la demanda 2000 - 2014	33
Gráfica 5. Datos discretizados de la generación 2000 - 2014	33
Gráfica 6. Datos discretizados de los aportes 2000 - 2014	34
Gráfica 7. Datos suavizados del precio de bolsa y su media	35
Gráfica 8. Datos suavizados de la demanda y su media	35
Gráfica 9. Datos suavizados de la generación y su media	36
Gráfica 10. Datos suavizados de los aportes y su media	37
Gráfica 11. Curva desviación estándar de los precios de bolsa	38
Gráfica 12. Curva desviación estándar de la demanda	39
Gráfica 13. Curva desviación estándar de la generación	40
Gráfica 14. Curva desviación estándar de los aportes	42
Gráfica 15. BagPlot de los precios de bolsa	43
Gráfica 16. BagPlot de la demanda	44
Gráfica 17. BagPlot de la generación	44
Gráfica 18. BagPlot de los aportes	45
Gráfica 19. Curva beta 0	47
Gráfica 20. Curva beta 1: demanda	47
Gráfica 21. Curva beta 2: generación	48
Gráfica 22. Curva beta 3: aportes	48
Gráfica 23. Datos suavizados del precio de bolsa	49
Gráfica 24. Datos suavizados del precio de bolsa estimado	49
Gráfica 25. Residuales y su media	50

## LISTA DE ECUACIONES

	pág.
Ecuación 1. Cálculo del precio de bolsa	20
Ecuación 2. Función de suavizado de curvas	21
Ecuación 3. Spline cúbica	23
Ecuación 4. Modelo lineal	27
Ecuación 5. Modelo en términos matriciales	27
Ecuación 6. Modelo concurrente	28
Ecuación 7. Modelo regresión múltiple concurrente	28
Ecuación 8. Modelo regresión múltiple concurrente notación tipo matriz	28

## LISTA DE ANEXOS

pág.

Anexo A. Estadísticos principales en los datos discretos resultados en R	59
--	----

## RESUMEN

En el mercado de energía mayorista de Colombia, participan como agentes los generadores quienes venden su energía a los usuarios finales mediante transacciones de contratos o en la bolsa de energía. Tanto las compras como las ventas en bolsa se transan a precio de bolsa, los contratos se establecen para mitigar la exposición a este precio por lo que se le puede considerar la variable más importante de la parte transaccional del mercado eléctrico. El precio de bolsa depende entre otras de decisiones de oferta, condiciones de disponibilidad de la generación, así como, de la incertidumbre en las expectativas de los generadores sobre el escenario hidrológico próximo o ante el efecto del fenómeno de El Niño o La Niña que inciden en el clima en nuestro país.

Con el presente estudio, se observa el comportamiento de las variables precio de bolsa, generación, aportes y demanda de energía en el periodo 2000 - 2014 a través de la observación de las mismas como datos funcionales, obteniendo la desviación y media funcional, medidas que permiten entender y analizar cada variable en un mismo momento del tiempo ( $t_j$ ) en los quince años comparados. Así mismo, permite identificar los periodos atípicos de las series funcionales para explicar en ellas las razones de su comportamiento a través del análisis del mercado. Por último, a través de un modelo de regresión funcional concurrente se determina si las variables generación, aportes y demanda de energía tienen efecto en la determinación del precio de bolsa nacional de energía.

Palabras clave:

- Mercado eléctrico
- Bolsa de energía
- Generación
- Aportes
- Demanda de energía
- Precio de bolsa
- Análisis de datos funcionales
- Regresión funcional concurrente.

## ABSTRACT

At the Colombian Electricity Market, generators participate as agents who sell energy to end users through contracts or transactions at the energy exchange market. Both purchases and sales are traded at the energy exchange market by energy spot price, contracts are set to mitigate exposure to this price so it can be considered the most important variable of the transactional component of the electricity market. The energy spot price depends among others on supply decisions, availability of generation's conditions, as well as uncertainty in the generators' expectations on the next hydrological scenario or with El Niño or La Niña phenomenon that have an effect on climate in our country.

In the present work, We study the behavior of spot price, generation, hydric contributions and energy demand variables between 2000 and 2014 by observing them through functional data analysis, obtaining the standard deviation of functional data and mean of functional data, useful measures for understanding and analyzing each variable in a specific moment of time ( $t_j$ ) in the fifteen years compared. It also identifies atypical periods of functional series in order to identify within them the reasons for their behavior through market analysis. Finally, through a concurrent functional linear model it determines whether the variables generation, hydric contributions and energy demand have an effect in determining the energy spot price.

Keywords:

- Electricity market
- Energy exchange market
- Generation
- Hydric contributions
- Energy demand
- Energy spot price
- Functional data analysis
- Concurrent functional linear model.

## 1. INTRODUCCIÓN

A partir de la Ley 142 y 143 del año 1994 - Régimen de Servicios Públicos Domiciliarios se privatiza las empresas de servicios públicos, entre ellos el servicio de energía, se entrega así la administración y operación del sistema eléctrico nacional a consorcios privados creándose un modelo de libre mercado de energía.

Colombia se reconoce como uno de los países con el mercado de energía mayorista más avanzado en América Latina; desde la creación en 1995 de este modelo, producto del racionamiento de 1992 (por el fenómeno de El Niño), no se ha vuelto a presentar en el sistema interconectado nacional - SIN, alguno que sea relevante en la demanda de energía del país.

En este mercado participan como agentes tanto generadores hidráulicos como térmicos, así como una pequeña parte de generación de energía renovable quienes venden su energía a los usuarios finales (\*) mediante transacciones de contratos o en la bolsa de energía.

Debido a la riqueza hídrica de Colombia, es de esta forma como se cubre la mayor parte de su demanda de energía eléctrica. La capacidad instalada efectiva neta de generación en Colombia a diciembre de 2014 es de aproximadamente 15,5 GW<sup>1</sup>. El 70.4 % de ella corresponde a generadores hidráulicos, 29.0 % a generadores térmicos y tan solo un 0.6 % se genera a partir de fuentes de energía renovable.

Cada generador participa en la bolsa diariamente con una oferta de precio y cantidad de energía. “El esquema definido para determinar el precio del mercado al cual se remuneran todas las transacciones de energía, es el despacho ideal, permite mediante un ordenamiento en mérito de las ofertas de las plantas, determinar cuál es la última planta con la que se cubre la atención de la demanda en cada una de las horas del día, para determinar el precio de bolsa horario”<sup>2</sup>.

---

<sup>1</sup> XM S.A. E.S.P. LOS EXPERTOS EN MERCADOS. Portal BI-Información Inteligente. Capacidad Efectiva Neta (kw) 2014. [En línea] Disponible en <<http://www.xm.com.co>>

(\*) Los usuarios finales se dividen en usuarios regulados (usuarios residenciales y pequeños usuarios con una capacidad instalada menor a 0.1 MW) y usuarios no regulados (grandes usuarios con una capacidad instalada mayor a 0.1 MW).

<sup>2</sup> ARENAS HOYOS, David. Mercado spot de energía y modelo alternativo para la fijación de un precio eficiente. Proyecto fin de carrera Profesional en Economía. Universidad EAFIT. Departamento de Economía. 2014. p. 4

Este trabajo es el resultado de buscar la respuesta a la pregunta ¿Cuál es el efecto de las variables generación, aportes y demanda de energía del sistema interconectado nacional (\*), en la determinación del precio de bolsa nacional de energía en Colombia para el periodo 2000 a 2014? Para resolver esta pregunta, se desea hacer uso de una metodología estadística que permita realizar dicha validación, lo que permitiría enfocar actuaciones en aquellas que demuestren tener impacto en el precio de bolsa nacional.

Este documento se estructura de la siguiente manera: después de esta introducción se realiza el planteamiento del problema, justificación y los objetivos, luego contiene una descripción del marco de referencia de la determinación del precio de bolsa de energía, seguido de la metodología. La sección 8 se refiere al análisis de los resultados encontrados. Finalmente, se establecen las conclusiones del estudio y la discusión final de los hallazgos en la formación del precio de bolsa de energía nacional.

## **2. PLANTEAMIENTO DEL PROBLEMA O PREGUNTA DE INVESTIGACIÓN**

¿Cuál es el efecto de las variables generación, aportes y demanda de energía del sistema interconectado nacional-SIN, en la determinación del precio de bolsa nacional de energía en Colombia para el periodo 2000 a 2014?

## **3. OBJETIVO GENERAL**

Determinar estadísticamente el efecto de las variables generación, aportes y demanda de energía del sistema interconectado nacional-SIN, en la determinación del precio de bolsa nacional de energía en Colombia para el periodo 2000 a 2014.

## **4. OBJETIVOS ESPECÍFICOS**

- Determinar la tendencia de los últimos quince años de dichas variables (generación<sup>3</sup>, aportes<sup>4</sup> y demanda de energía del sistema

---

(\*) Demanda del SIN = generación + importaciones + demanda no atendida - exportaciones.

<sup>3</sup> XM S.A. E.S.P. LOS EXPERTOS EN MERCADOS. Portal BI-Información Inteligente. Generación 2000-2014. [En línea] Disponible en <<http://www.xm.com.co>>

<sup>4</sup> XM S.A. E.S.P. LOS EXPERTOS EN MERCADOS. Portal BI-Información Inteligente. Aportes Diario 2000-2014. [En línea] Disponible en <<http://www.xm.com.co>>

interconectado nacional<sup>5</sup>), así como del precio de bolsa nacional de energía<sup>6</sup> a través de la observación de las mismas como datos funcionales.

- Identificar los periodos atípicos de las series funcionales en los periodos analizados (2000 a 2014).
- Analizar estadísticamente e identificar en los periodos atípicos las razones de su comportamiento a través del análisis fundamental.
- Determinar si las variables generación, aportes y demanda de energía del sistema interconectado nacional tienen efecto en la determinación del precio de bolsa nacional de energía.

## 5. JUSTIFICACIÓN

El precio de bolsa se puede considerar la variable más importante de la parte transaccional del mercado eléctrico. Tanto las compras como las ventas en bolsa se transan al precio de bolsa, los contratos son un mecanismo de cubrimiento ante la incertidumbre del precio de bolsa, la escasez de generación del sistema se estima a partir de la diferencia entre precio de bolsa y el precio de escasez, y así existen muchas otras señales de mercado relacionadas con este precio.

El precio de bolsa depende principalmente de las decisiones de oferta y las condiciones de disponibilidad de la generación hidráulica, en este mercado eléctrico constituido de forma mayoritaria por este tipo de generación. Ahora bien, el comportamiento creciente del precio de bolsa durante las transiciones de invierno y verano refleja entre otros aspectos, la incertidumbre en las expectativas de los generadores sobre el escenario hidrológico próximo, expresadas a través de su oferta.

Así mismo, entre los factores que inciden en la formación del precio de bolsa en el territorio colombiano se destaca el comportamiento de las variables climáticas y oceánicas en el Pacífico tropical. Esto se debe a que las perturbaciones periódicas generadas por cualquiera de los extremos de las variables asociadas con El Niño

---

<sup>5</sup> XM S.A. E.S.P. LOS EXPERTOS EN MERCADOS. Portal BI-Información Inteligente. Demanda Energía SIN 2000-2014. [En línea] Disponible en <<http://www.xm.com.co>>

<sup>6</sup> XM S.A. E.S.P. LOS EXPERTOS EN MERCADOS. Portal BI-Información Inteligente. Precio Bolsa Nacional SIN 2000-2014. [En línea] Disponible en <<http://www.xm.com.co>>



y la Oscilación del Sur (ya sea El Niño o La Niña), modulan el comportamiento del clima en nuestro país y la distribución de las precipitaciones a nivel estacional.

Es importante realizar un estudio en este sentido, teniendo en cuenta que los altos precios de energía, afectan tanto la economía de los residenciales como la competitividad de los industriales. Validar estadísticamente el efecto de las variables observadas en la formación del precio de bolsa, puede permitir enfocar actuaciones en aquellas que más impacto tienen en el mismo.

## **6. MARCO DE REFERENCIA**

Para comprender el interés en resolver esta pregunta de investigación es necesario conocer el sector eléctrico en Colombia, cómo está estructurado el mercado de energía, cómo se transa la energía y los diferentes entes y actores que participan en este sector.

### **6.1 ESTRUCTURA DEL MERCADO<sup>7</sup>**

El sector eléctrico se fundamenta en el hecho de que las empresas comercializadoras y los grandes consumidores adquieren la energía y potencia en un mercado de grandes bloques de energía, que opera libremente de acuerdo con las condiciones de oferta y demanda. El precio de la electricidad en este mercado se establece de común acuerdo entre las partes contratantes, sin la intervención del Estado.

En Colombia, la empresa XM Compañía de Expertos en Mercados S.A. ESP opera el sistema interconectado nacional (SIN) en la que se coordina la operación de la cadena productiva del sector eléctrico colombiano a mediano y largo plazo, a través de la planeación los recursos de generación y los recursos de transmisión, de acuerdo con la demanda de energía eléctrica actual y en el corto plazo, realizando la recepción de las ofertas diarias que presentan los generadores en la Bolsa de Energía, donde se asignan hora a hora las plantas que suministrarán la energía al día siguiente; y administra el mercado de energía mayorista colombiano (MEM) en la que atiende las transacciones comerciales, liquida y factura los intercambios de energía resultantes entre los agentes generadores y comercializadores del mercado, recauda el dinero producto de las transacciones

---

<sup>7</sup> XM S.A. E.S.P. LOS EXPERTOS EN MERCADOS. Informe de Operación del Sistema y Administración del Mercado Eléctrico Colombiano 2000-2014. [En línea] Disponible en <<http://www.xm.com.co>>

en bolsa, y las transacciones internacionales de electricidad (TIE), así como, el recaudo de los servicios por transmisión nacional y regional para entregarlos a los agentes transmisores y distribuidores por el uso de sus redes.

## **6.2 ORGANISMOS DE ADMINISTRACIÓN Y CONTROL DEL MERCADO ELÉCTRICO COLOMBIANO**

El Régimen de servicios públicos está fundamentado en la Ley 142 de 1994 y en particular por la Ley 143 del mismo año para el servicio público de electricidad en sus actividades definitorias de mercado. Los organismos encargados de la correcta operación, evolución y control del sector de energía eléctrica en Colombia se listan a continuación:

- Ministerio de Minas y Energía
- Unidad de Planeación Minero Energética UPME
- Superintendencia de Servicios Públicos SSP
- Comisión de Regulación de Energía y Gas CREG
- Centro Nacional de Despacho CND
- Consejo Nacional de Operación CNO

## **6.3 MERCADO ELÉCTRICO COLOMBIANO<sup>8</sup>**

**6.3.1 Operaciones comerciales.** El mercado colombiano tiene dos grandes integrantes: el mercado spot y el mercado de contratos, asociándose el primero al corto plazo y el segundo al largo plazo.

Dentro del mercado de contratos se realizan operaciones de compraventa de energía a través de un contrato como instrumento financiero. Los contratos pueden suscribirse entre generadores, entre comercializadores, entre generadores y comercializadores y entre comercializadores y usuarios finales.

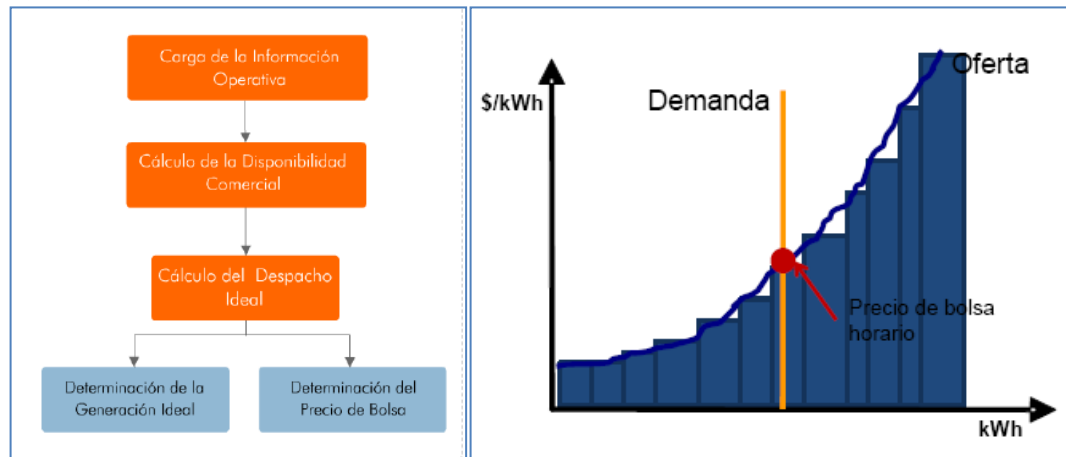
En el mercado spot se liquidan todas aquellas transacciones que no tienen ámbito dentro del mercado de contratos. En el mercado colombiano se cuentan las siguientes operaciones: compra - venta de energía en bolsa, reconciliaciones, venta de servicio AGC (Automatic Generation Control), pago responsabilidad comercial servicio AGC, cargo por confiabilidad, restricciones, desviaciones, transacciones internacionales de energía.

---

<sup>8</sup> EMGESA S.A. E.S.P. Documento Conceptual de Generación 2014.

**6.3.2 Despacho ideal.** El proceso de despacho ideal toma la demanda comercial que es igual en magnitud a la generación real y por tanto cubre la demanda real y las pérdidas para cruzarla contra la disponibilidad comercial de los generadores con el fin de asignarles generación a las plantas y determinar el precio de bolsa (ver figura 1), junto a otras variables necesarias para la liquidación pero sin tener en cuenta ningún tipo de restricción eléctrica u operativa del sistema eléctrico.

Figura 1. Flujo de información para el despacho ideal – Determinación del precio de bolsa.



Fuente: XM, S.A. E.S.P. Memorias de capacitaciones y seminarios 2007-2015. [En línea] Disponible en <<http://www.xm.com.co>>.

## 6.4 DEFINICIONES<sup>9</sup>

**6.4.1 Precio de bolsa nacional.** En condiciones normales de operación, corresponde al mayor precio de oferta de las unidades con despacho centralizado que han sido programadas para generar en el despacho ideal y que no presentan inflexibilidad. Representa un precio único para el sistema interconectado en cada período horario.

En condiciones de intervención de precios de oferta, se determina de acuerdo con el procedimiento para condiciones normales de operación, pero teniendo en cuenta los precios intervenidos de oferta para las plantas de generación hidroeléctrica con embalse, definidos en el Código de Operación.

<sup>9</sup> XM S.A. E.S.P. LOS EXPERTOS EN MERCADOS. Glosario. 2015. [En línea] Disponible en <<http://www.xm.com.co>>

**6.4.1.1 Cálculo del precio de bolsa nacional.** El cálculo del precio de bolsa nacional para un periodo horario está dado por la siguiente ecuación:

Ecuación 1. Cálculo del precio de bolsa

$$\text{Precio de Bolsa Nacional}_{\text{Periodo } x} = \text{MPO}_{\text{Nacional}_{\text{Periodo } x}} + \Delta I_{\text{Mercados Nacionales día}}$$

Fuente: CREG. Resolución 160 de 2009

Donde,

MPO: máximos precios ofertados

$\Delta I$ : delta de incremento diario para mercados nacionales en pesos por kWh

**6.4.2 Demanda.** Cantidad de energía consumida por un agente comercializador o generador, medida en sus fronteras comerciales.

**6.4.2.1 Demanda de energía.** Carga en kWh solicitada a la fuente de suministro en el punto de recepción durante un período de tiempo determinado.

**6.4.2.2 Demanda total doméstica.** Sumatoria de los valores de la demanda doméstica de todos los comercializadores, que incluye los factores de pérdidas para referir a nivel de 220 kV y las pérdidas del STN.

**6.4.3 Generación.** Producción de energía eléctrica, generalmente en una central hidráulica o térmica.

**6.4.3.1 Generación real.** Se calcula como la sumatoria de las generaciones netas medidas a nivel horario para cada uno de los agentes generadores en sus puntos de frontera.

**6.4.3.2 Generación neta.** Generación entregada por una planta al Sistema Interconectado Nacional en el punto de conexión.

## **6.5 ANÁLISIS DE DATOS FUNCIONALES (FDA)**

Las variables funcionales se caracterizan por la evolución de una variable a lo largo del tiempo (proceso estocástico), de modo que los valores que toman son funciones.

El análisis de datos funcionales transforma los datos discretos en funciones; la unidad básica de información es la función completa, a la que se le aplica diversas

técnicas con el objetivo de descubrir estructuras que con el análisis de datos discretos no sería posible evidenciar.

El análisis de datos funcionales es aquella parte de la estadística que trabaja con muestras de funciones aleatorias. Ramsay y Silverman (1997), entre otros, han introducido herramientas de análisis de para este tipo de datos. Las medidas de tendencia central, de dispersión y de relación entre variables aplicadas a muestras de variables aleatorias se pueden definir de manera análoga para muestras de datos funcionales. Sólo hace falta considerar que ahora estaremos trabajando en un espacio vectorial distinto: el espacio  $L_2$  (\*) (funciones de cuadrado integrable). En realidad Ramsay y Silverman definen estas medidas manera general para cualquier espacio que tenga definido un producto escalar, y luego deduce de ellas las definiciones para muestras de funciones<sup>10</sup>.

**6.5.1 Suavizado de funciones mediante funciones B-splines<sup>11</sup>.** El suavizado de curvas a través de funciones consiste en aproximar las funciones del espacio considerado a través de:

Ecuación 2. Función de suavizado de curvas

$$\chi(t) = \sum_{l=1}^K c_l B_l(t)$$

Fuente: RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. Second Edition. United States of America: Springer Science + Business Media. Inc. 2005, p. 44.

donde  $K$  es el número de funciones de la base.

Una base es un conjunto de funciones conocidas e independientes  $\phi_k, k \in N$  tales que cualquier función puede ser aproximada, tan bien como se quiera, mediante una combinación lineal  $K$  de ellas. Las bases más usadas para datos funcionales

(\*) Sea  $T = [a, b] \subset R$ ,  $L_2(T) = \{ f \rightarrow R, \text{tal que } \int_t^x f(t)^2 dt < \infty \}$

<sup>10</sup> RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. 1997, citado por NAVARRO PÉREZ, Valentín. Análisis de datos funcionales: implementación y aplicaciones. Proyecto fin de carrera Licenciatura en Ciencias y Técnicas Estadísticas. Universidad Politécnica de Catalunya. Facultad de Matemáticas y Estadística. Julio 2004. p. 7

<sup>11</sup> RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. Second Edition. United States of America: Springer Science + Business Media. Inc. 2005, p. 43-52

son las bases de Fourier si los datos son periódicos, bases B-spline para cálculos rápidos y flexibles, bases de Wavelets apropiada para discontinuidades, entre otras.

Una base Fourier es especialmente útil para las funciones extremadamente estables, es decir, las funciones de las que no hay características locales fuertes y donde la curvatura tiende a ser del mismo orden en todas partes. Idealmente, la periodicidad de la base Fourier debe reflejarse en cierto grado en los datos.

Generalmente producen expansiones que son suaves y uniformes, pero no son apropiadas en cierto grado para los datos conocidos o sospechosos de reflejar discontinuidades en la función en sí o en derivadas de orden inferior. Estas características han hecho de la base Fourier la base de tradicional elección para grandes series de tiempo, pero nuevas técnicas como la B-splines y wavelets puede igualar e incluso superar esta eficiencia computacional.

Las funciones spline son la opción más común de sistema de aproximación para datos funcionales o parámetros no periódicos. Los splines combinan el cálculo rápido de los polinomios con sustancialmente mayor flexibilidad, a menudo alcanzado con sólo un pequeño número de funciones de base. Además, los sistemas básicos que se han desarrollado para funciones spline requieren una cantidad de cálculo que es proporcional a  $n$ , una propiedad importante, ya que muchas aplicaciones implican miles o millones de observaciones.

Aunque hay muchas maneras de que estos sistemas se pueden construir, el sistema de base B-spline es el más popular, y el código está disponible en una amplia gama de lenguajes de programación, incluyendo R, con el que se realizó este trabajo.

La propiedad de que una función de base B-splines de orden  $m$  es positivo durante no más de  $m$  intervalos, y que éstos son adyacentes, se llama la propiedad de soporte compacto, y es de la mayor importancia para el cálculo eficiente computacional.

Si hay  $k$  funciones de base B-spline, entonces la matriz de orden  $K$  de los productos interiores de estas funciones será la banda estructurada, con sólo 1  $m$ -sub-diagonales por encima y por debajo de la diagonal principal que contiene valores distintos de cero. Esto significa que no importa lo grande que es  $K$ , y que se ocupe de los valores en los miles, el cálculo de la función spline puede ser

organizada con el fin de aumentar linealmente con la única  $K$ . De esta manera, las bases B-splines comparten potencialmente las ventajas computacionales de los sistemas de bases ortogonales como las bases Fourier y bases wavelet.

Otra propiedad de la base B-splines es la característica de tiempo o espacialmente-localizada que se refiere a las opciones en cuanto a donde deben colocarse los breakpoints interiores o knots. Muchas aplicaciones por defecto aplican una separación de igual distancia, lo cual está bien, siempre y cuando los datos son relativamente equidistantes. Si no es así, puede ser más prudente colocar un knot en cada punto de datos  $j$ -enésimo, dónde  $j$  es un número fijado de antemano. Esto equivale a la colocación de knots interiores en los cuantiles de la distribución. Un caso especial es el de splines suavizantes, donde un breakpoint se coloca en cada valor del argumento. Por último, se puede optar por ignorar cualquiera de estos simples procedimientos para colocar más knots en las regiones que se sabe contienen alta curvatura, y menos donde hay menos.

Por características expuestas de la base B-splines de eficiencia computacional, de ser útil para cálculos rápidos y flexibles con miles de observaciones, y teniendo en cuenta las características de los datos recogidos de las variables en las que se pudo observar desde los datos discretizados que la curvatura no es del mismo orden en todas partes, y que los datos son equidistantes, se opta por la utilización de esta base para el desarrollo de este trabajo.

Una spline cúbica es una función  $S$  definida sobre  $T$ , tal que  $S$  es un polinomio cúbico en el intervalo  $[T_{l-1}, T_l]$ ,  $l = 1, 2, 3, \dots, L + 1$  con segunda derivada continua, dado un conjunto de  $L$  puntos interiores del intervalo  $T = [a, b]$ , es decir,  $a = T_1 < T_2 < \dots < T_L = b$ . Entre mayor sea el número de los puntos  $T_l$ , denominados nodos, la curva presentará mayor flexibilidad. Una spline cúbica  $S$  con nodos  $T_0 < T_1 < \dots < T_L$ , se expresa como:

Ecuación 3. Spline cúbica

$$S(t) = \sum_{k=0}^{L+4} c_k B_k(t)$$

Fuente: RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. Second Edition. United States of America: Springer Science + Business Media. Inc. 2005, p. 51.

donde  $c$  es un vector  $k$ -dimensional de coeficientes y  $B(t)$  es una función  $L + 4$ -dimensional con componentes  $B_k(t)$ .

**6.5.2 Análisis descriptivo en datos funcionales<sup>12</sup>.** Como ya se mencionó las medidas de tendencia central, de dispersión y de relación entre variables aplicadas a muestras de variables aleatorias se aplican igualmente para muestras de datos funcionales, sólo que en este caso, los objetos calculados corresponden a curvas.

Las correspondientes funciones descriptivas de la media y la desviación estándar utilizadas en este trabajo están dadas por las expresiones:

Dado un conjunto de datos funcionales  $\chi_1, \chi_1, \dots, \chi_n$  definidos en  $t \in T \in \mathbb{R}$ ,

- Media:  $\bar{\chi}(t) = \frac{\sum_{i=1}^n \chi_i(t)}{n}$
- Varianza:  $Var(\chi(t)) = \frac{\sum_{i=1}^n \chi_i(t) - (\bar{\chi}(t))^2}{n-1}$
- Desviación estándar:  $\sqrt{Var(\chi(t))}$

**6.5.3 Herramientas gráficas del análisis de datos funcionales<sup>13</sup>.** Existen actualmente diversas herramientas para visualizar grandes cantidades de datos funcionales en forma de curvas suavizadas; dentro de ellas se incluyen el gráfico arco iris, el diagrama de caja funcional, y el diagrama de región de mayor densidad (HDR) funcional, que sirven como métodos de análisis a través de una interfaz gráfica.

Estos métodos facilitan el hallazgo de características que podrían no haber sido evidentes con el uso de estadísticos y modelos matemáticos, como lo es la identificación con precisión de valores atípicos con gran velocidad de cálculo, mientras que proporciona una representación gráfica, que pueden no ser identificadas a partir de un gráfico de los datos originales.

El diagrama de caja funcional y el diagrama HDR funcional identifican valores atípicos en los datos funcionales. El diagrama de caja funcional muestra la curva media (la curva con la mayor profundidad), y las regiones interior y exterior. La región interior se define como la región limitada por todas las curvas correspondientes a puntos en la bolsa bivariada. Por lo tanto, 50% de las curvas

<sup>12</sup> RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. Second Edition. United States of America: Springer Science + Business Media. Inc. 2005, p. 22-25

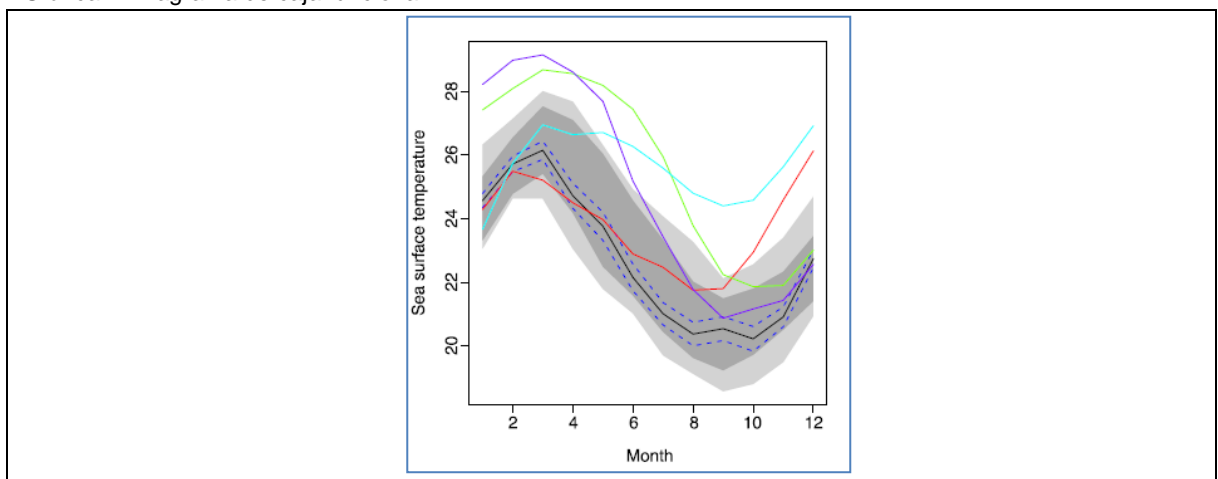
<sup>13</sup> HYNDMAN, Rob J., & SHANG, Han Lin. Rainbow Plots, Bagplots, and Boxplots for Functional Data. Journal of Computational and Graphical Statistics. 2010, 19:1, 29-45. [En línea] Disponible en <<http://dx.doi.org/10.1198/jcgs.2009.08158>>



están en la región interior. La región exterior se define de manera similar como la región limitada por todas las curvas correspondientes a los puntos dentro de la región límite exterior bivariada.

Cuando los valores atípicos están lejos de la mediana, el diagrama de caja funcional se considera un buen método de detección de valores atípicos, sin embargo, si los valores extremos se encuentran cerca de la mediana, puede identificarlos incorrectamente, en este caso, el diagrama de caja HDR funcional es considerado más apropiado.

Gráfica 1. Diagrama de caja funcional.



Fuente: HYNDMAN, Rob J., & SHANG, Han Lin. Rainbow Plots, Bagplots, and Boxplots for Functional Data. *Journal of Computational and Graphical Statistics*. 2010, 19:1, 29-45. [En línea] Disponible en <<http://dx.doi.org/10.1198/jcgs.2009.08158>>.

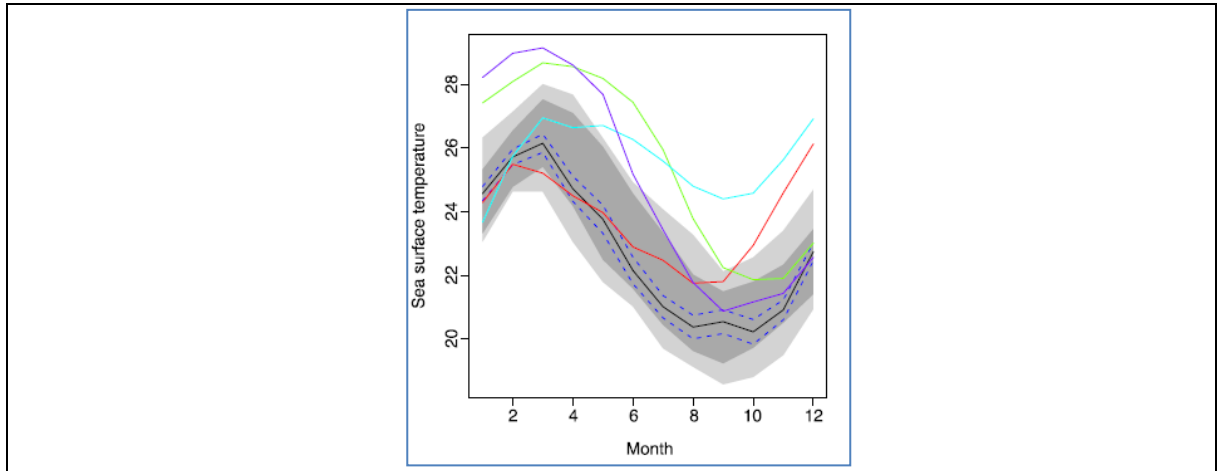
En la gráfica 1 se observa un diagrama de caja funcional, las líneas negras son las curvas de la mediana, rodeado de intervalos de confianza del 95%. Las curvas fuera de las regiones límite o cerco se muestran como valores atípicos en diferentes colores.

El diagrama de caja HDR funcional se basa en el diagrama de caja HDR bivariado (Hyndman 1.996), utilizando las estimaciones de densidad kernel, un HDR es definido como  $R\alpha = \{z : \hat{f}(z) \geq f\alpha\}$ , donde  $f\alpha$  es tal que  $\int_{R\alpha} \hat{f}(z) dz = 1 - \alpha$ ; es decir, es la región con una cobertura de probabilidad  $1 - \alpha$ , donde todos los puntos dentro de la región tienen una estimación de densidad más alta que cualquiera de los puntos de fuera de la región, de ahí el nombre de "región de mayor densidad." Para una densidad bivariada, las regiones de más alta densidad

pueden ser consideradas como contornos, con una ampliación de cobertura de  $\alpha$  disminuciones.

El diagrama de caja HDR funcional muestra la curva modal (la curva con la densidad más alta), y las regiones interior y exterior. La región interior se define como la región limitada por todas las curvas correspondientes a los puntos dentro del 50% del HDR bivariado. Por lo tanto, 50% de las curvas están en la región interior. La región exterior se define de manera similar como la región limitada por todas las curvas correspondientes a los puntos dentro del borde exterior del HDR bivariado.

Gráfica 2. Diagrama HDR funcional.



Fuente: HYNDMAN, Rob J., & SHANG, Han Lin. Rainbow Plots, Bagplots, and Boxplots for Functional Data. Journal of Computational and Graphical Statistics. 2010, 19:1, 29-45. [En línea] Disponible en <<http://dx.doi.org/10.1198/jcgs.2009.08158>>.

En la gráfica 2 se observa un diagrama HDR funcional, las líneas negras son las curvas modales. Las curvas fuera de la HDR exterior se muestran como valores atípicos en diferentes colores.

**6.5.4 Modelo de regresión lineal funcional** <sup>14</sup>. El modelo de regresión funcional es la extensión natural del modelo de regresión lineal que permite establecer la relación entre una variable respuesta y una o más covariables, con la diferencia que existe por lo menos una variable (respuesta o explicativa) funcional en dicho modelo.

<sup>14</sup> RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. Second Edition. United States of America: Springer Science + Business Media. Inc. 2005, p. 217-221, 247-260

De acuerdo a la propuesta de Ramsay and Silverman (2005) asumiendo que se tienen  $M$  tratamientos cada uno con un número  $n_m$  de sujetos y que  $y(t)$  es una respuesta funcional, el modelo lineal para la  $i$ -ésima función en el  $g$ -ésimo grupo ( $y_{ig}(t)$ ), está dado por:

Ecuación 4. Modelo lineal

$$y_{ig}(t) = \mu(t) + \alpha_g(t) + \varepsilon_{ig}(t)$$

Fuente: RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. Second Edition. United States of America: Springer Science + Business Media. Inc. 2005, p. 218.

donde la función  $\mu(t)$  es la media general,  $\alpha_g(t)$  representa la función media para cada tratamiento y  $\varepsilon_{ig}(t)$  es la función de error en cada caso. Para determinar el modelo, se debe establecer la matriz diseño con el fin de estimar los parámetros funcionales  $\mu(t)$  y  $\alpha_g(t)$ , cumpliendo la condición de que  $\sum_g \alpha_g(t) = 0$  para todo  $t$  con el fin de asegurar la estimabilidad de los mismos.

En términos matriciales el modelo puede expresarse como:

Ecuación 5. Modelo en términos matriciales

$$y_{ig}(t) = \sum_{j=1}^{G+1} x_{(ig)j} \beta_j(t) + \varepsilon_{(ig)}(t)$$

$$Y(t) = \chi(t) \beta(t) + \varepsilon(t)$$

Fuente: RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. Second Edition. United States of America: Springer Science + Business Media. Inc. 2005, p. 255.

La matriz diseño  $\chi$  tiene la misma estructura que en el caso de regresión múltiple o simple, la diferencia es que el vector de parámetros  $\beta(t)$  y las predicciones  $X\beta(t)$ , son vectores de funciones en vez de vectores de números.

Para determinar cómo una o más variables explicativas influyen en la variable respuesta específicamente en el tiempo, existen varios modelos de acuerdo a la respuesta que estamos buscando, concurrente, anual o total, influencia local, entre otros. Presentamos los conceptos asociados al modelo de regresión lineal concurrente usado en la aplicación. Una revisión detallada puede hacerse en Ramsay and Silverman, 2005.

Considerando un modelo cuya variable respuesta  $Y$  depende sólo de una variable explicativa  $\chi$  se puede utilizar  $\chi$  en el tiempo  $s = t$ . El modelo es:

Ecuación 6. Modelo concurrente

$$Y_i(t) = \alpha(t) + \chi_i(t)\beta(t) + \varepsilon_i(t)$$

Fuente: RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. Second Edition. United States of America: Springer Science + Business Media. Inc. 2005, p. 220.

Este modelo es llamado concurrente o puntual.

Al considerar un modelo con  $q$  covariables funcionales  $z_{ij}$  cada una multiplicada por su coeficiente de regresión funcional  $\beta_j$  el modelo de regresión múltiple concurrente es:

Ecuación 7. Modelo de regresión múltiple concurrente

$$Y_i(t) = \sum_{j=1}^q z_{ij}(t)\beta_j(t) + \varepsilon_i(t)$$

Fuente: RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. Second Edition. United States of America: Springer Science + Business Media. Inc. 2005, p. 255.

Si  $N$  por  $q$  matriz funcional  $Z$  contiene esas  $z_{ij}$  y el vector del coeficiente de regresión funcional  $\beta$  de tamaño  $q$  contiene cada una de las funciones de regresión, el modelo de regresión múltiple concurrente en notación tipo matriz es:

Ecuación 8. Modelo de regresión múltiple concurrente notación tipo matriz

$$y(t) = Z(t)\beta(t) + \varepsilon(t)$$

Fuente: RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. Second Edition. United States of America: Springer Science + Business Media. Inc. 2005, p. 255.

Donde  $Y$  es un vector funcional de tamaño  $N$  que contienen las funciones respuesta.

## 7. METODOLOGÍA

### 7.1 TIPO DE ESTUDIO

Este es un estudio de tipo descriptivo dado que tiene como finalidad describir la tendencia, identificar, y analizar estadísticamente los periodos atípicos de las series funcionales, observados en un periodo de quince años (2000 a 2014) de las variables escogidas, así como identificar y validar si existe efecto de las variables

explicativas (generación, aportes y demanda de energía) en la variable respuesta (precio de bolsa nacional de energía).

## 7.2 MÉTODO

En este caso en el que los datos que se estudian son procesos continuos (generación de energía, cotización bursátil del precio de bolsa, etc.) se utiliza un método estadístico; específicamente una especialidad estadística conocida con el nombre de Análisis de Datos Funcionales (FDA).

## 7.3 VARIABLES DE ESTUDIO Y UNIDADES DE ANÁLISIS

Se recolectó y organizó los datos históricos diarios de los últimos quince años de las variables, teniendo las siguientes consideraciones:

- **Precio de bolsa nacional \$/kWh:** las cifras originales descargadas del portal BI de XM S.A. E.S.P. Expertos en Mercados disponibles en línea vienen horarias, en pesos por kilovatio hora, 365 días año, por lo que se realiza un promedio aritmético diario, con el fin de estandarizar todas las variables a cifras diarias. Así mismo, se elimina el dato correspondiente al día 29 de febrero de los años bisiestos (2000, 2004, 2008 y 2012) para que las series sean todas de 365 días en los quince años observados, es decir, se eliminan cuatro datos conservando 5.475 por lo que se consideran que no tenga un impacto relevante en el análisis final.
- **Demanda de energía SIN GWh:** las cifras originales descargadas del portal BI de XM S.A. E.S.P. Expertos en Mercados disponibles en línea vienen en kilovatios hora, diarias, 365 días año, se convierten a gigavatios hora para el manejo de cifras más pequeñas. Al igual que la variable precio de bolsa nacional, se elimina el dato correspondiente al día 29 de febrero de los años bisiestos (2000, 2004, 2008 y 2012) para que las series sean todas de 365 días en los quince años observados, es decir, se eliminan cuatro datos conservando 5.475 por lo que se consideran que no tenga un impacto relevante en el análisis final.
- **Generación SIN GWh:** las cifras originales descargadas del portal BI de XM S.A. E.S.P. Expertos en Mercados disponibles en línea vienen en kilovatios hora, diarias, 365 días año, se convierten a gigavatios hora para el manejo

de cifras más pequeñas. Al igual que las variables anteriores, se elimina el dato correspondiente al día 29 de febrero de los años bisiestos (2000, 2004, 2008 y 2012) para que las series sean todas de 365 días en los quince años observados, es decir, se eliminan cuatro datos conservando 5.475 por lo que se consideran que no tenga un impacto relevante en el análisis final.

- **Aportes de energía SIN GWh:** las cifras originales descargadas del portal BI de XM S.A. E.S.P. Expertos en Mercados disponibles en línea, vienen por región y por río, en kilovatios hora, 365 días año, por lo que se realiza una sumatoria diaria, con el fin de estandarizar todas las variables a cifras diarias, adicionalmente se convierten a gigavatios hora para el manejo de cifras más pequeñas. Así mismo, se elimina el dato correspondiente al día 29 de febrero de los años bisiestos (2000, 2004, 2008 y 2012) para que las series sean todas de 365 días en los quince años observados, es decir, se eliminan cuatro datos conservando 5.475 por lo que se consideran que no tenga un impacto relevante en el análisis final.

## 7.4 PROCEDIMIENTO

A juicio del investigador se decidió incluir para el estudio una variable desde la oferta: generación de energía, una variable desde la demanda: demanda del sistema y finalmente una variable correspondiente a la hidrología: aportes del sistema interconectado nacional, teniendo en cuenta el aspecto ya mencionado de la incidencia en la formación del precio de bolsa nacional por el comportamiento de las variables climáticas.

La realización de este estudio se hizo a través de tres fases: obtención de la Información, depuración de datos y análisis de la información que se describen a continuación.

**7.4.1 Fase 1. Obtención de la información.** Se recolectaron los datos históricos de los últimos quince años de las variables generación, demanda y aportes de energía del sistema interconectado nacional-SIN, así como del precio de bolsa nacional, obtenidos del portal de información BI de la página de internet de XM Compañía de Expertos en Mercados, que es la empresa que opera y administra el mercado eléctrico colombiano.

**7.4.2 Fase 2. Depuración de datos.** Teniendo en cuenta que la información obtenida originalmente viene desagregada de manera diferente para cada variable y con el fin de estandarizar todas las variables a cifras diarias, se realizó para cada una de ellas el tratamiento puntal que fue necesario... véase el numeral 7.3...

Así mismo, dado el método estadístico que se utilizó de análisis de datos funcionales se elimina el dato correspondiente al día 29 de febrero de los años bisiestos (2000, 2004, 2008 y 2012) para que las series sean todas de 365 días en los quince años observados.

**7.4.3 Fase 3. Análisis de la información.** Para el análisis de los datos se ha utilizado el software de licencia gratuita R<sup>15</sup>, que es un lenguaje de alto nivel y un entorno para el análisis de datos y gráficos. Puede descargarse libremente del sitio llamado CRAN (Comprehensive R Archive Network) en la siguiente dirección: <http://cran.r-project.org/>. Existen tres librerías de R sobre FDA con las que se trabaja en el presente estudio que se encuentran disponibles en CRAN:

- fda versión 2.4.4 (Autores: J. O. Ramsay, Hadley Wickham, Spencer Graves, Giles Hooker)
- fda.usc versión 1.2.1 (Autores: Manuel Febrero-Bande, Manuel Oviedo de la Fuente)
- rainbow versión 3.3 (Autores: Han Lin Shang, Rob J Hyndman)

## **7.5 ASPECTOS ÉTICOS**

El siguiente es un extracto de los términos legales uso PORTAL BI<sup>16</sup> propiedad de XM S.A. E.S.P. en lo que corresponde al uso de la información pública la que ha sido utilizada en el presente trabajo.

El servicio de PORTAL BI de la compañía XM S.A. E.S.P publica información pública y confidencial respecto de la operación, la administración y el desarrollo del Mercado Eléctrico Mayorista en Colombia.

PORTAL BI permite al visitante el acceso a la información pública y provee el sistema de claves para el acceso a la información privada y confidencial. El

---

<sup>15</sup> R CORE TEAM 2015. Version 3.2.2. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [En línea] Disponible en <<https://www.R-project.org/>>

<sup>16</sup> XM S.A. E.S.P. LOS EXPERTOS EN MERCADOS. Términos legales uso Portal BI XM. 2014. [En línea] Disponible en <<http://www.xm.com.co>>

acceso y uso de este sitio web implican aceptación de su parte de los términos y condiciones legales de la misma.

XM pone a disposición del usuario (persona natural o jurídica que utiliza el servicio de PORTAL BI), de forma gratuita y libre de cualquier cargo, información de carácter pública, a través de su portal.

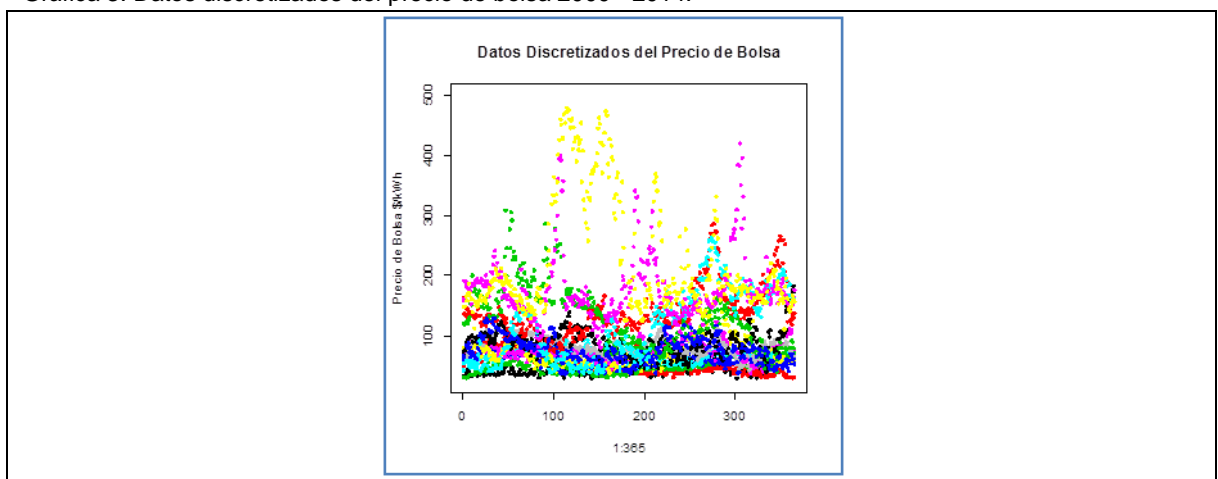
Dicha información está normada por la política publicada en su página web y son en parte la reproducción de información suministrada por terceros y en parte información procesada, elaborada y desarrollada por XM. Sobre esta última parte; existen derechos de autor a favor de XM. Por lo anterior, solo se autoriza reproducción de información procesada, elaborada y desarrollada por XM con cita de la fuente es decir: consultado en el portal de XM S.A. E.S.P. y con indicación de la fecha de la consulta.

## 8. RESULTADOS

### 8.1 DATOS DISCRETIZADOS

Luego de crear las matrices en R con las observaciones de la base de datos, se pueden observar en las gráficas 3, 4, 5 y 6 que se encuentran a continuación, los datos discretizados de las variables precio de bolsa nacional en \$/kWh, demanda de energía SIN en GWh, generación SIN en GWh y aportes de energía SIN en GWh.

Gráfica 3. Datos discretizados del precio de bolsa 2000 - 2014.

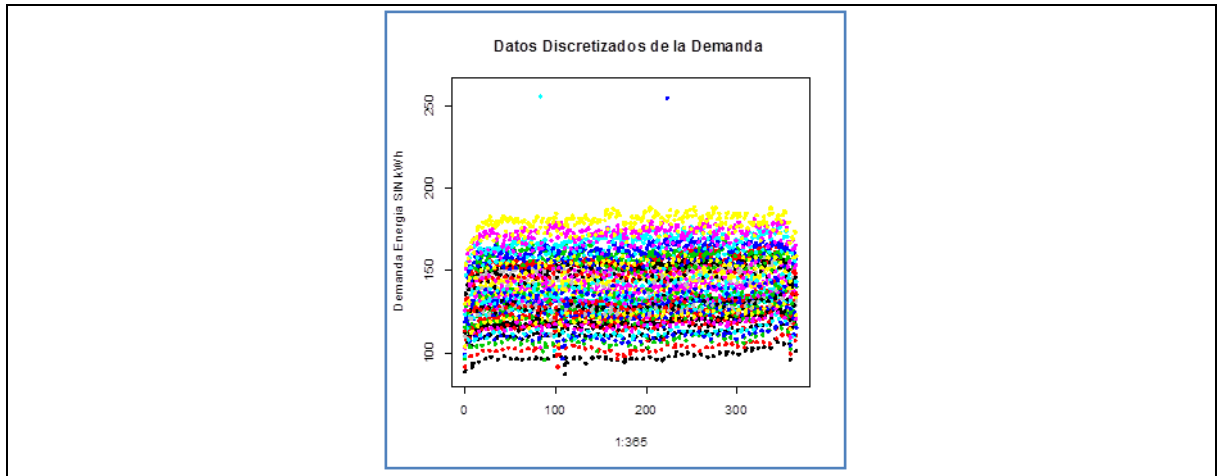


Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.



Tal como lo podemos observar en la gráfica 3, son pocos los días en el que el precio promedio diario supera el rango de los 400 \$/kWh; sólo 40 días de los 5475 datos considerados (0.7%) presenta este comportamiento, todos en el primer semestre del año 2014, en los meses de abril, mayo y junio, a excepción del 1ro. de noviembre del 2013 con 420.40 \$/kWh.

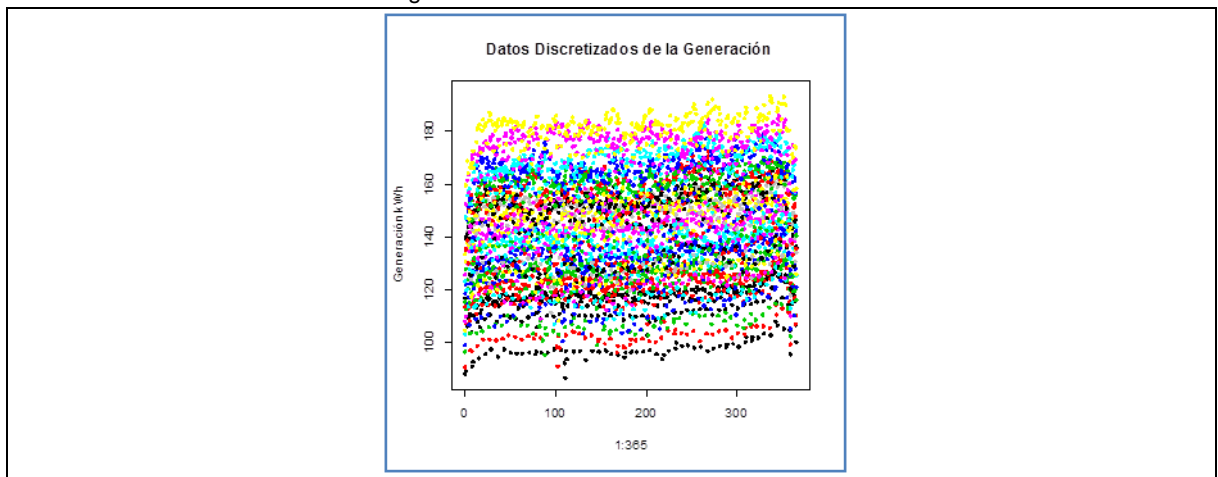
Gráfica 4. Datos discretizados de la demanda 2000 - 2014.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R

Para el caso de la demanda tal como lo podemos observar en la gráfica 4, los datos se mueven en el rango de 86,79 y 188,15 GWh siendo superados sólo por 254,49 y 255,39 GWh correspondientes al 11 de agosto de 2011 y el 25 de marzo de 2012 respectivamente, puntos fácilmente observables en la gráfica.

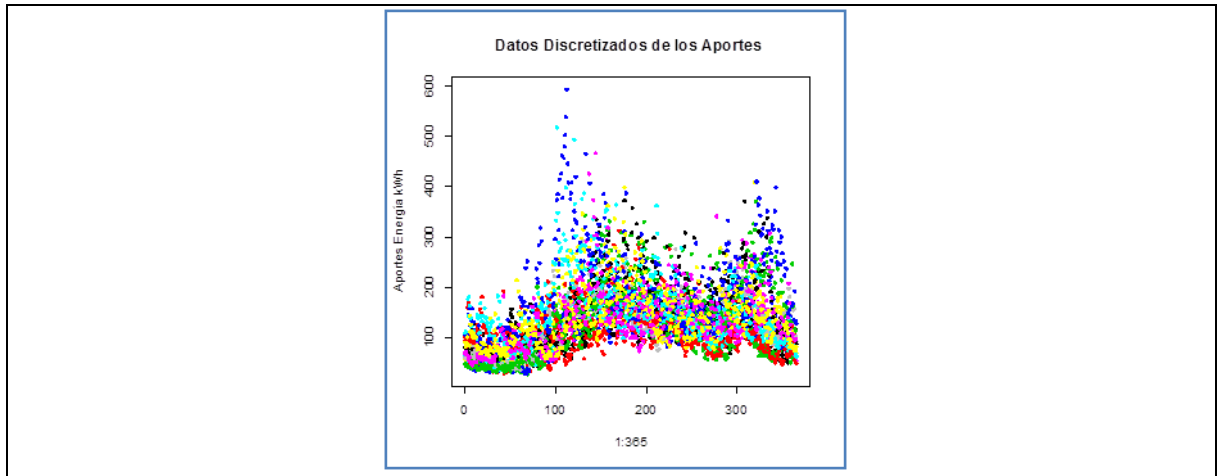
Gráfica 5. Datos discretizados de la generación 2000 - 2014.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

La generación tiene un comportamiento aún más compacto que la demanda observable en la gráfica 5, los datos se mueven entre 86,53 y 193,36 GWh sin que se presente ninguna cifra fuera de ese rango.

Gráfica 6. Datos discretizados de los aportes 2000 - 2014.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

La variable aportes por su parte como se observa en la gráfica 6, presenta datos bastante dispersos porque su rango es amplio: entre 25,25 y 591,25 GWh lo que es de entender teniendo en cuenta que dicha variable mide el aporte en GWh de las fuentes hídricas del sistema, hallándose sólo el 2,2% de los mismos por encima de 300 GWh.

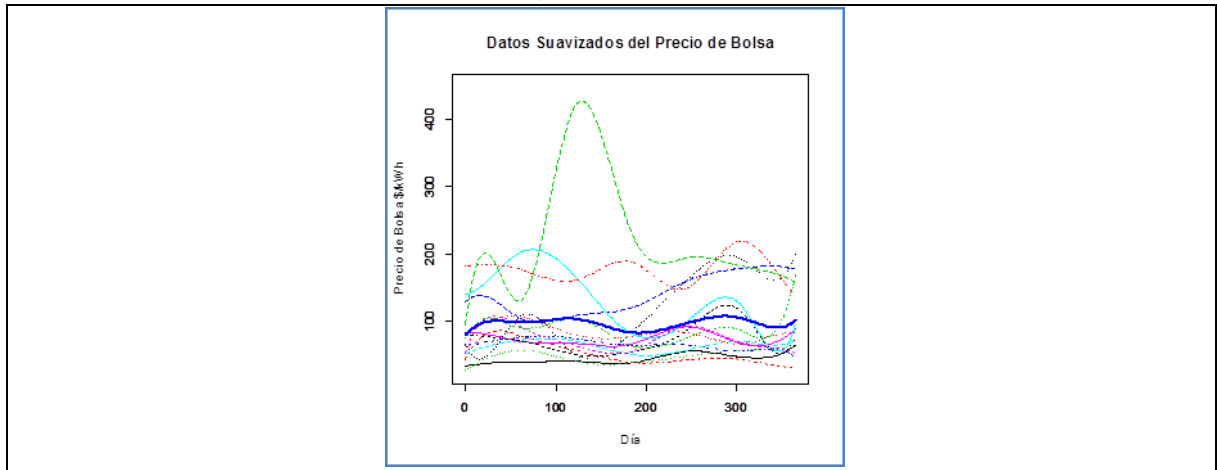
## 8.2 DATOS SUAVIZADOS Y MEDIA FUNCIONAL

A través de la función `Data2fd` se convierten las matrices generadas anteriormente en un objeto de datos funcionales realizándose así la suavización de datos observable en las gráficas 7, 8, 9 y 10 que se encuentran a continuación y que corresponden a los datos suavizados de las mismas variables.

Como fue citado en el marco de referencia ... véase el numeral 6.5 ..., las medidas de tendencia central, de dispersión y de relación entre variables se pueden definir también para muestras de datos funcionales; es justamente lo que nos permite realizar R al aplicar la función `Data2fd`, creando así objetos de datos funcionales a los cuales se les puede calcular en este caso la media y la desviación estándar con las funciones `mean.fd` y `sd.fd` respectivamente.

La función `mean.fd` evalúa la media del conjunto de funciones en cada objeto de datos funcionales definido para cada variable observada; en cada gráfico, se puede identificar como la curva azul de mayor grosor.

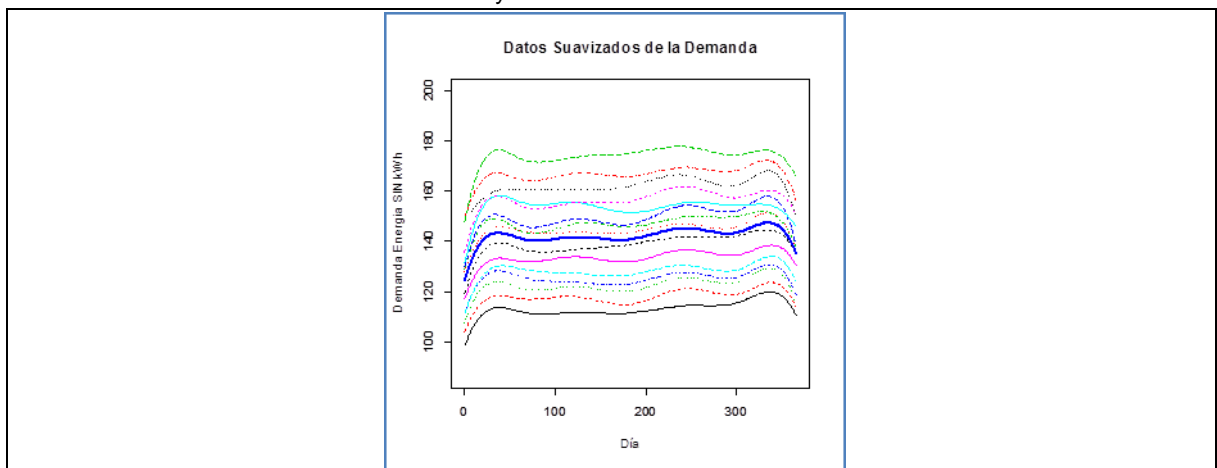
Gráfica 7. Datos suavizados del precio de bolsa y su media.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

Es interesante observar que a pesar del comportamiento alto del precio promedio de bolsa en el 2014, que se puede ver en la gráfica 7 como la línea punteada verde y como fue mencionado en la observación de los datos discretizados de dicha variable ... véase el numeral 8.1 ... este no tiene un impacto más considerable en la media funcional; es así como dicha medida del precio medio diario de bolsa de energía oscila entre 80.00 y 120.00 \$/kWh aproximadamente (curva azul de mayor grosor, en el gráfico), en los 365 días del año.

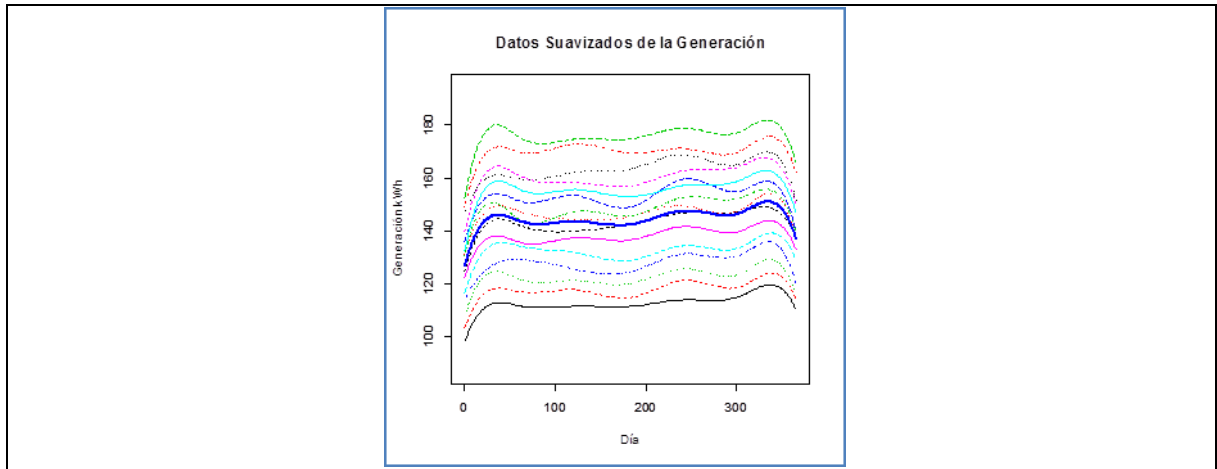
Gráfica 8. Datos suavizados de la demanda y su media.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

Todas las series de la variable demanda tienen un comportamiento similar observable en la gráfica 8 en el que cae los últimos días del año se empieza a recuperar en los primeros días del año, manteniéndose estable el resto del tiempo, por consiguiente, la media funcional tiene esa misma tendencia estando la mayor parte del tiempo alrededor de 140 GWh en promedio (curva azul de mayor grosor, en el gráfico), y siendo menor en los últimos y primeros días del año.

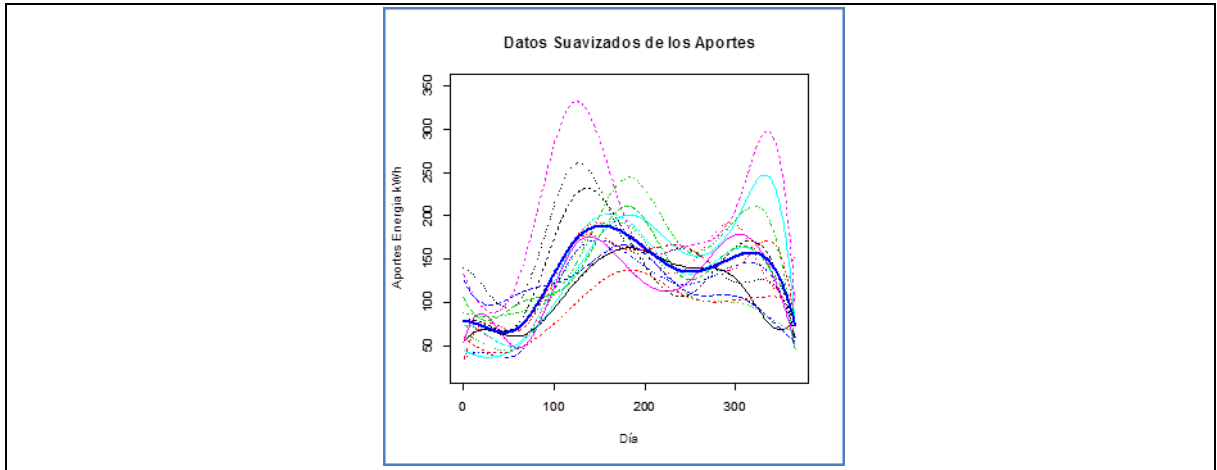
Gráfica 9. Datos suavizados de la generación y su media.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

La variable generación tiene un comportamiento muy similar al de la variable demanda observable en la gráfica 9, lo que es de suponer teniendo en cuenta que en el mercado de la energía la oferta y la demanda deben coincidir en el tiempo debido a que la energía es un bien no almacenable, es así como, la media funcional tiene esa misma tendencia estando la mayor parte del tiempo ligeramente por encima de los 140 GWh en promedio (curva azul de mayor grosor, en el gráfico), y siendo menor en los últimos y primeros días del año.

Gráfica 10. Datos suavizados de los aportes y su media.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

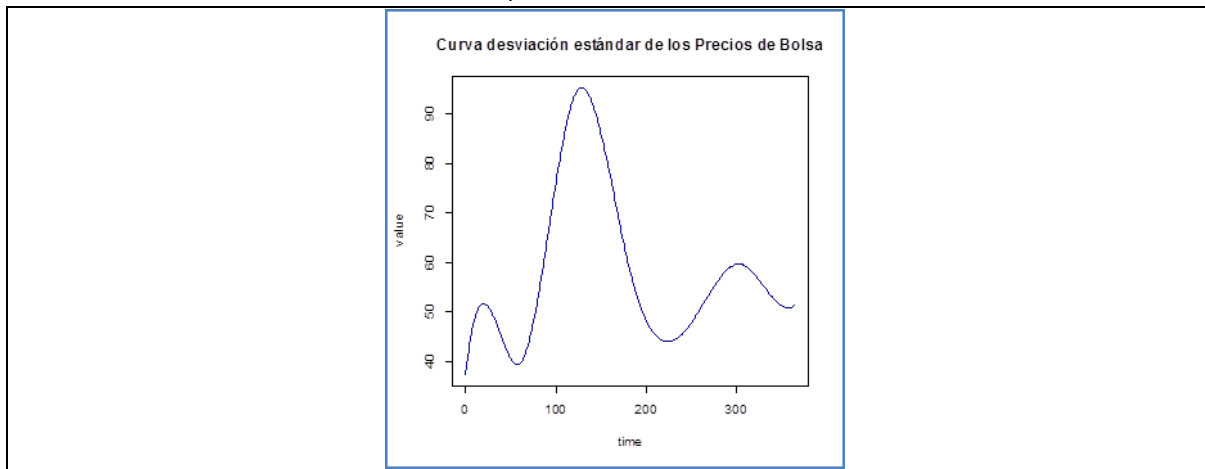
Tal como fue mencionado en la observación de los datos discretizados ... véase el numeral 8.1 ... la variable aportes es la que mayor fluctuación presenta observable en la gráfica 10, lo que se refleja en su media funcional que tiene esa misma tendencia moviéndose en un rango aproximado entre 60 y 190 GWh en promedio (curva azul de mayor grosor, en el gráfico), presentando sus valores máximos entre los días 140 y 160 del año aproximadamente que corresponden a los últimos diez días de mayo y los primeros diez días de junio, con un nuevo incremento de menor intensidad entre los días 275 y 325 del año aproximadamente que corresponden a los meses de octubre y noviembre consideradas épocas de lluvias en Colombia.

Al igual que en las variables demanda y generación los valores más pequeños se presentan en los últimos y primeros días del año pero la recuperación al inicio del año de esta variable es más aplanada que las dos anteriores comenzando su incremento alrededor de los días 50 y 60 del año que corresponde a la última semana de febrero y una caída más fuerte que dichas variables finalizando el año a partir del día 325 aproximado que corresponde a la última semana de noviembre hasta finalizar el año lo que es consistente con lo que se considera la temporada seca que en Colombia correspondiente a los meses de diciembre y enero aunque esto puede variar bastante por la presencia de los fenómenos de El Niño y La Niña ya antes mencionados.

### 8.3 DESVIACIÓN ESTÁNDAR

La función `sd.fd` evalúa la desviación estándar del conjunto de funciones en cada objeto de datos funcionales definido para cada variable observada; las gráficas 11, 12, 13 y 14 corresponden a los gráficos de dicha medida de dispersión para cada variable.

Gráfica 11. Curva desviación estándar de los precios de bolsa.



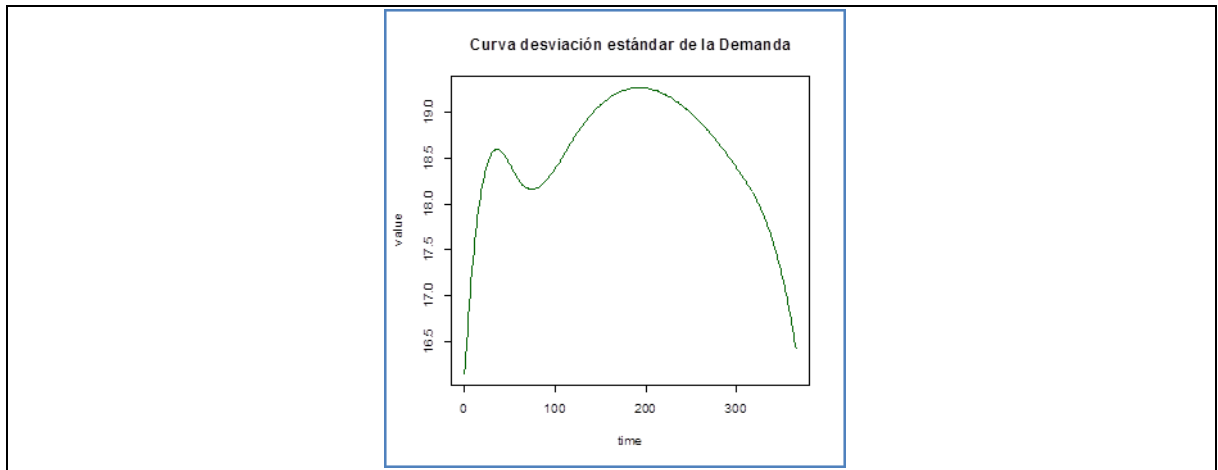
Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

La desviación estándar para el caso del precio de bolsa tal como lo podemos ver en la gráfica 11, nos muestra que los precios varían más alrededor del día 125 que corresponde a los primeros días de mayo con dos puntos más de menor variación, la primera semana de octubre y los últimos días de enero. Las épocas de menor variación se dan los primeros días del año, la última semana de febrero seguido de los últimos días de septiembre y los primeros de octubre.

Tal como lo menciona Alejandro Lucio Gerente General de Derivex S.A. en el 2011 el “precio de bolsa es variable y altamente volátil, pues depende de varios fundamentales como las condiciones climáticas actuales y sus perspectivas futuras, el precio de commodities a su vez con altos niveles de volatilidad como gas, carbón y otros combustibles, el crecimiento económico, etc.”<sup>17</sup>, así como el efecto de fenómenos como El Niño y La Niña, la disponibilidad de la generación, entre otros.

<sup>17</sup> LUCIO, Alejandro. ¿Es posible reducir los costos de energía eléctrica con derivados financieros? Revista Dinero [En línea] (30 de agosto de 2011 7:40:00 AM) disponible en <<http://www.dinero.com/inversionistas/articulo/es-posible-reducir-costos-energia-electrica-derivados-financieros/134196>>

Gráfica 12. Curva desviación estándar de la demanda.

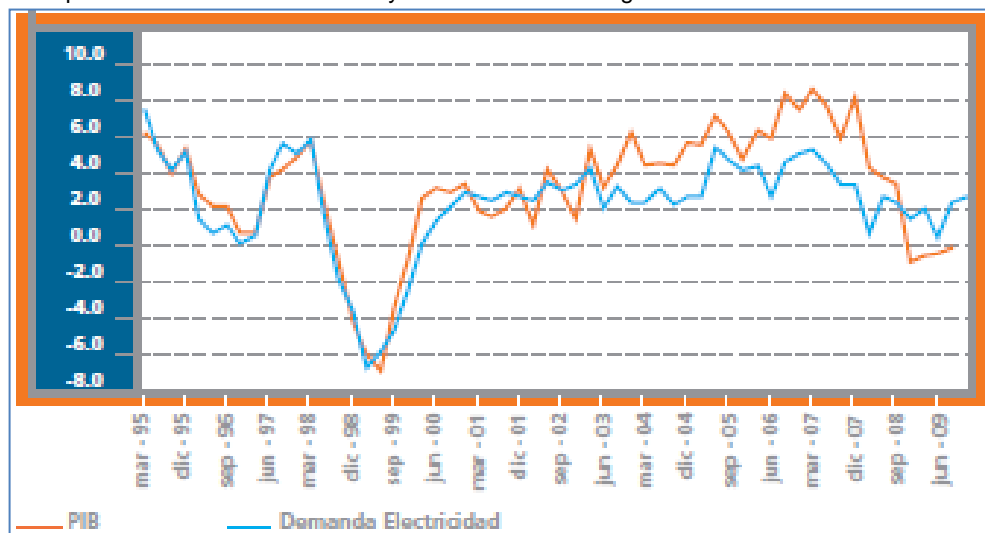


Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

En la demanda como se observa en la gráfica 12, la desviación estándar indica que varía más entre finales de junio y primeros días de agosto con punto de menor variación en la última semana de enero y primera de febrero. Las épocas de menor variación se dan los primeros y últimos días del año.

Son varios los factores que explican las variaciones que presenta la demanda de energía, de hecho, su comportamiento es diferente por tipo de día (comerciales, sábado y domingo-festivos).

Figura 2. Comportamiento trimestral del PIB y la demanda de energía



Fuente: XM, S.A. E.S.P. Informe de Operación del Sistema y Administración del Mercado 2011. [En línea] Disponible en <<http://www.xm.com.co>>.

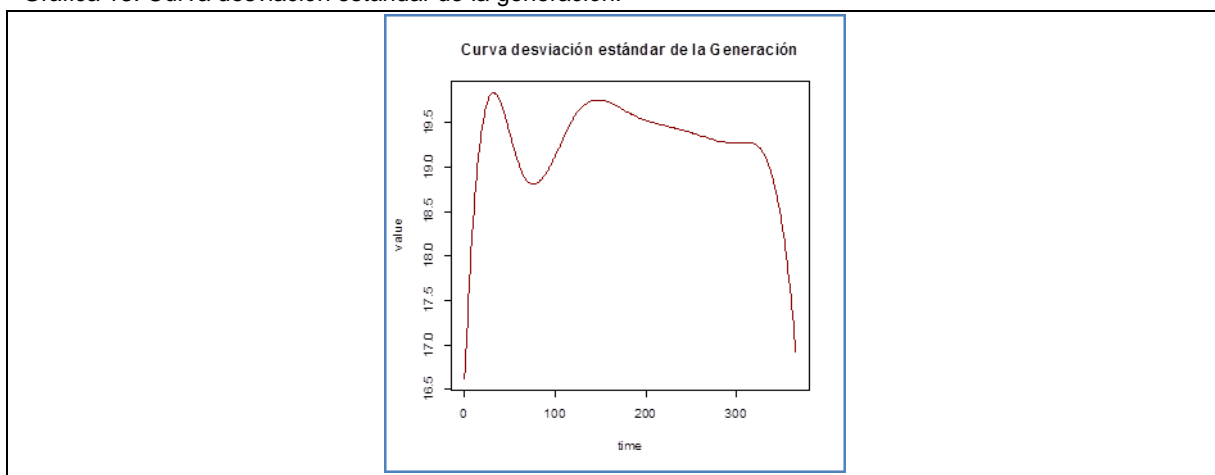
El comportamiento del Producto Interno Bruto-PIB del país es una de las variables que más relación tiene con la demanda de energía, tal como se muestra en la figura 2 y por lo tanto su variación, aunque se considere una inelasticidad parcial de la demanda, debido a que los hogares mantienen un consumo mínimo de energía (mercado regulado) a pesar del impacto que pueda ocasionar en el mercado no regulado una crisis económica a escala global como la del 2008-2009.

Una de las principales razones de la variación la demanda del mercado regulado (consumos residenciales y pequeños negocios), es la ocasionada por el consumo de energía en refrigeración y acondicionamiento ante la presencia de las altas temperaturas; es así como la temporada de verano su duración y potencia explica dichas variaciones, así como a la presencia de El Niño por las altas temperaturas (mayor consumo) y por La Niña por las bajas temperaturas (menor consumo).

Entre tanto la demanda de energía del mercado no regulado que atiende las empresas clasificadas en las diferentes actividades económicas (industria manufacturera, explotación de minas y canteras, comercio, transporte, agropecuario, construcción, entre otras) varía por el crecimiento o contracción de las mismas.

En este mismo mercado la ocurrencia de un evento (huelgas, mantenimientos y voladuras a oleoductos) de algunas empresas como Cerromatoso, Cerrejón, Occidental entre otras, las cuales debido a su consumo representan el 5% de la demanda total del país, también afectan el comportamiento de la demanda.

Gráfica 13. Curva desviación estándar de la generación.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.



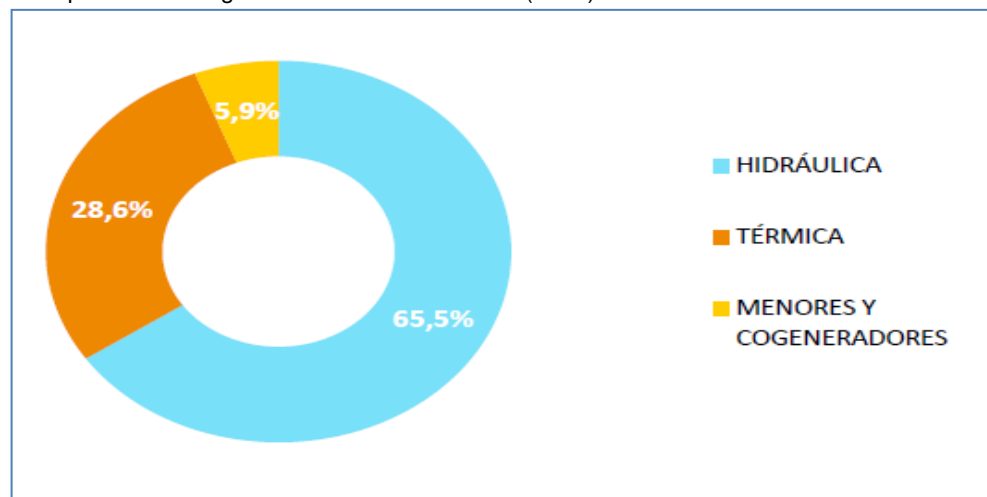
De similar comportamiento que la demanda la desviación estándar de la variable generación observable en la gráfica 13, indica que su mayor variación se presenta la última semana de enero con seguido de una menor variación en casi todo el mes de mayo, con un comportamiento más plano de variación similar entre inicios de junio y mediados de noviembre. Al igual que en la demanda las épocas de menor variación se dan los primeros y últimos días del año.

Debido a la naturaleza de este mercado en el que la oferta y la demanda deben coincidir en el tiempo por la imposibilidad de almacenar la energía los motivos que explican la variación de la demanda interna como las exportaciones de energía explican la variación de la generación.

Adicionalmente como se observa en la figura 3, la alta participación de la generación hidráulica en la composición de la generación del SIN hacen a esta variable muy susceptible al comportamiento climático en el país.

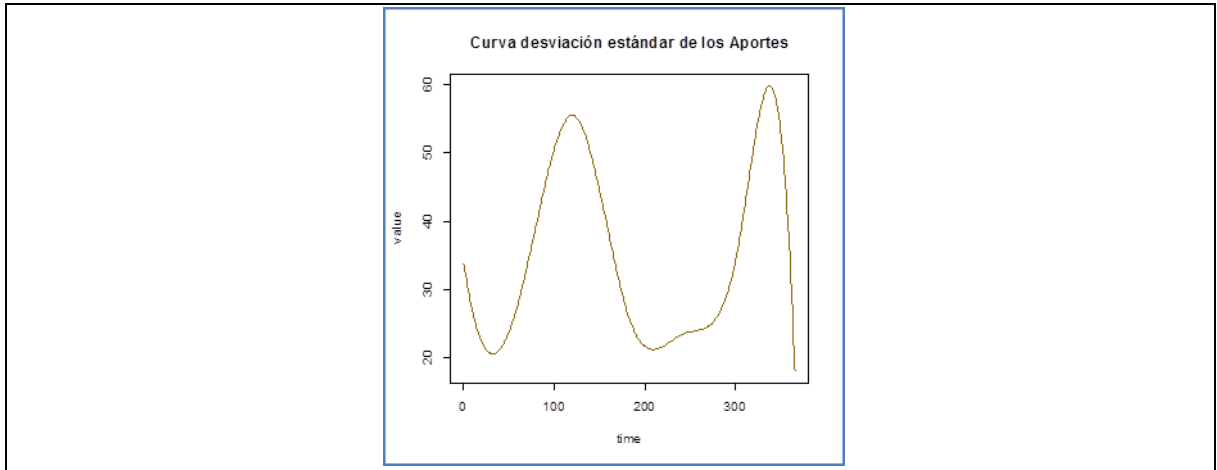
Por su parte la generación térmica se ve afectada por el precio y disponibilidad de commodities también de gran volatilidad como carbón, gas y otros combustibles requeridos por este tipo de generación.

Figura 3. Composición de la generación del SIN en 2014 (GWh)



Fuente: XM, S.A. E.S.P. Informe de Operación del Sistema y Administración del Mercado 2014. [En línea]  
Disponibile en <<http://www.xm.com.co>>.

Gráfica 14. Curva desviación estándar de los aportes.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

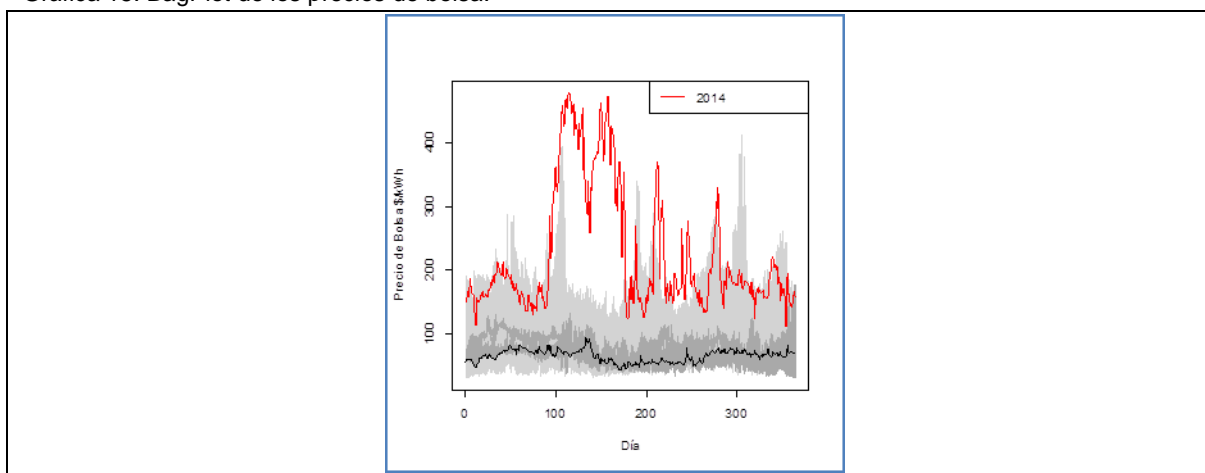
Como lo podemos ver en la gráfica 14, para el caso de los aportes la desviación estándar nos muestra que varían más alrededor del día 340 que corresponde a los primeros días de diciembre seguido, en menor variación, durante los primeros días de mayo. Las épocas de menor variación se dan los últimos días del año, los últimos días de enero seguido de los últimos diez días de julio y la primera semana de agosto.

Por la naturaleza de esta variable la presencia de los fenómenos de El Niño y La Niña tienen fuerte influencia en su variación así como el comportamiento de la época de lluvias y la temporada seca en Colombia.

#### 8.4 IDENTIFICACIÓN DE SERIES ATÍPICAS

Las gráficas BagPlot 15, 16, 17 y 18 nos permiten identificar los años que presentan comportamientos atípicos para cada variable.

Gráfica 15. BagPlot de los precios de bolsa.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

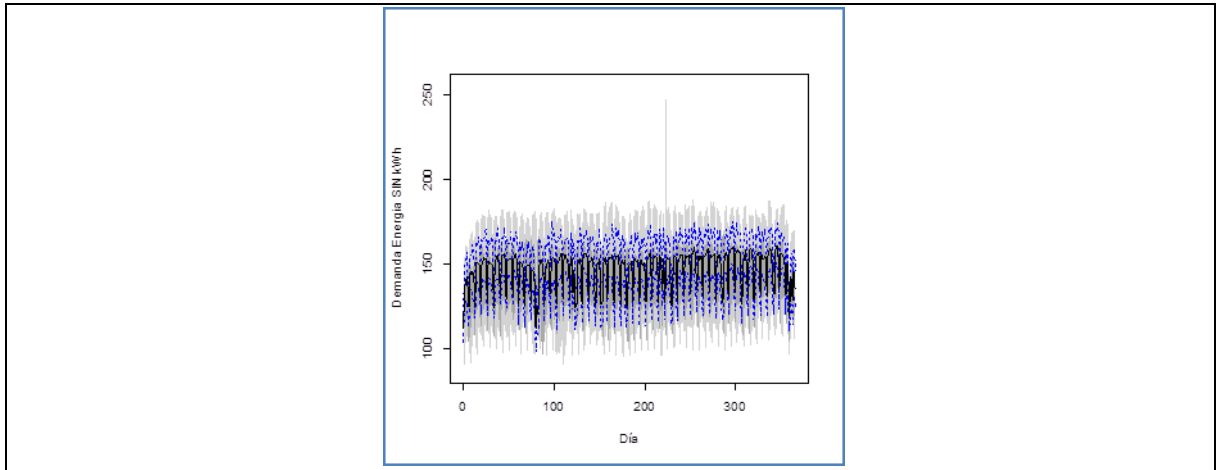
Tal como podemos observar en la gráfica 15, correspondiente al precio de bolsa de los 15 años observados del periodo 2000-2014 es el año 2014 el que presenta un comportamiento atípico para esta variable.

De acuerdo a XM operador del mercado<sup>18</sup> el precio de bolsa nacional en 2014 presentó un incremento anual del 26.07%, al pasar de un promedio ponderado anual en 2013 de 178.88 \$/kWh a 225.51 \$/kWh en 2014, con un máximo valor del precio de bolsa nacional horario el día 30 de abril de 480.54 \$/kWh y el valor mínimo el día 01 de enero con un valor de 38.94 \$/kWh.

Durante el 2014, el precio de bolsa superó el precio de escasez en el mes de abril los días comprendidos entre el 24 y 30, en el mes de mayo los días 01, 02, 05, 10 y 13, en el mes de junio los días 05, 06, 07, 09, 10, 11 y en el mes de agosto el día 01; de acuerdo al informe de este ente los altos precios de bolsa observados en abril, mayo y junio de 2014 se deben a que en estos meses se presentó poca disponibilidad del parque de generación hidráulico y a la gran oleada de calor que se presentó en estos meses.

<sup>18</sup> XM S.A. E.S.P. LOS EXPERTOS EN MERCADOS. Informe de operación del sistema y administración del mercado. 2014. [En línea] Disponible en <<http://www.xm.com.co>>

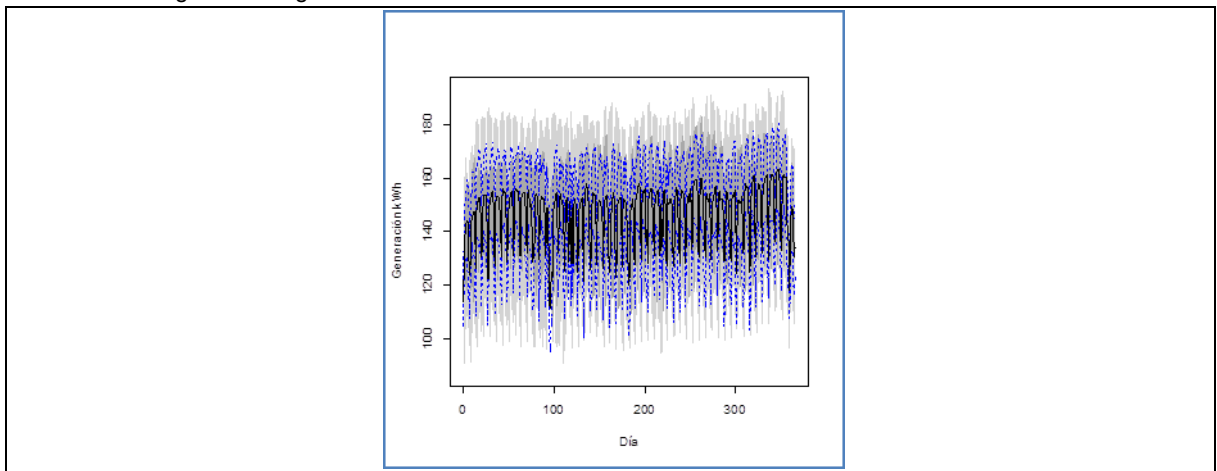
Gráfica 16. BagPlot de la demanda.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

Como podemos observar en esta gráfica 16, no se identifica ningún año de los 15 años observados del periodo 2000 -2014 que presente un comportamiento atípico para la variable demanda.

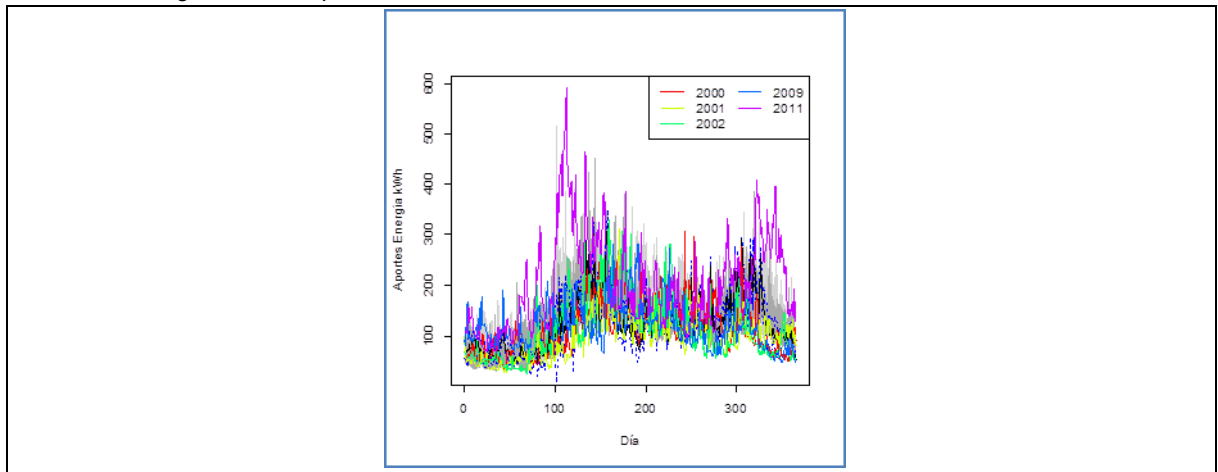
Gráfica 17. BagPlot de la generación.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

De la misma manera, de acuerdo a lo observable en la gráfica 16, no se identifica comportamiento atípico en ninguno de los 15 años observados del periodo 2000 - 2014 para la variable generación, gráfico 17.

Gráfica 18. BagPlot de los aportes.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

A diferencia de las anteriores, para la variable aportes como se observa en esta gráfica 18, los años 2000, 2001, 2002, 2009 y 2011 presentan comportamientos atípicos.

Al inicio de la serie en el año 2000 los aportes energéticos de origen hídrico estuvieron por encima del promedio histórico durante ocho de los doce meses del año a excepción de abril, junio, noviembre (a nivel del promedio) y diciembre seguido de aportes por debajo de los promedios históricos en diez de los doce meses del año, para los años 2001, 2002 y en nueve de los doce meses del año 2003 efecto de el fenómeno de El Niño 2002-2003.

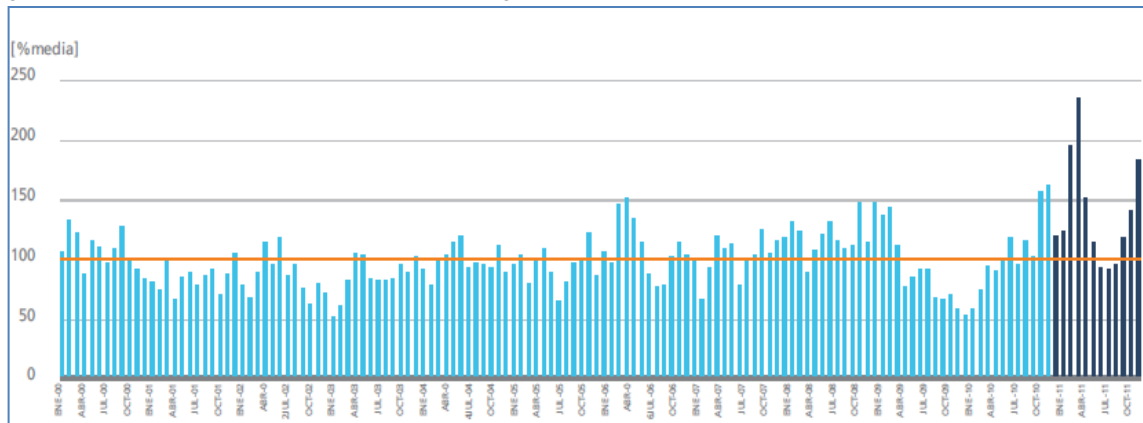
El comportamiento de los aportes para el 2009 de acuerdo a los análisis de XM operador del mercado es acorde con el efecto del fenómeno de El Niño 2009-2010, al finalizar dicho año los ríos asociados al SIN aportaron un total de 43,139.0 GWh, equivalentes al 88.8% de la media histórica (48,591.3 GWh), valor por debajo del registrado en el año 2008, cuando los aportes alcanzaron el 119% de la media. La ocurrencia continua durante ocho meses (mayo - diciembre 2009) de aportes por debajo de los promedios históricos, superó en número de meses a la registrada durante El Niño 2002-2003.

El comportamiento de los aportes para el 2011 de acuerdo a los análisis del operador, se debe principalmente a que estos fueron fuertemente influenciados por lo que se conoce como fenómeno de La Niña (FLN), extremo climático, que se asocia con la fase fría de un fenómeno físico que hace parte del ciclo natural

global del clima. Lo intenso de este fenómeno en el 2011 ocasionó precipitaciones excesivas sobre la mayor parte del territorio nacional.

En la figura 4 se muestra la evolución de aportes energéticos (hidráulicos) al SIN, del 2000 al 2011, expresados en porcentaje de la media. La línea horizontal naranja, corresponde a los aportes promedio (100%). Se observa, aportes por debajo de los promedios históricos, por el impacto del fenómeno de El Niño 2009-2010, que supera en número de meses a la registrada durante El Niño 2002-2003, seguido de La Niña 2010-2011 con sus condiciones típicas, durante los últimos meses del año 2011.

Figura 4. Evolución histórica de los aportes energéticos al SIN 2000-2011.



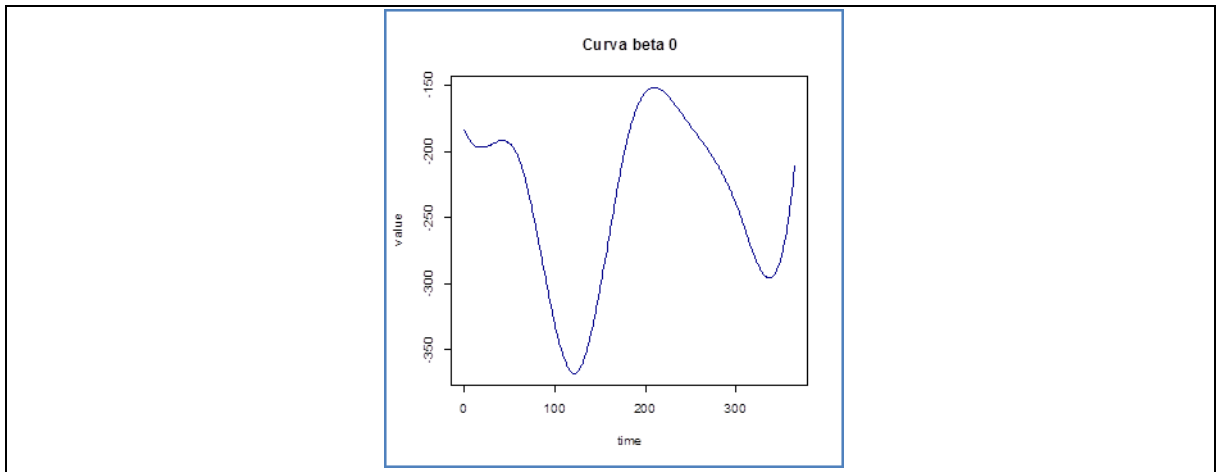
Fuente: XM, S.A. E.S.P. Informe de Operación del Sistema y Administración del Mercado 2011. [En línea]  
Disponible en <<http://www.xm.com.co>>.

## 8.5 REGRESIÓN FUNCIONAL CONCURRENTE

La función de R fRegress lleva a cabo un análisis de regresión funcional, donde ya sea la variable dependiente o de una o más variables independientes son funcionales. Las variables no funcionales pueden utilizarse a ambos lados de la ecuación. En el modelo lineal funcional concurrente para una variable dependiente funcional, todas las variables funcionales están evaluados en un momento en común o valor del argumento  $t$ . Es decir, el ajuste se define en términos del comportamiento de todas las variables en un momento en el tiempo determinado.

Las gráficas 19, 20, 21 y 22 corresponden a los gráficos de las curvas  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$  y  $\beta_3$ .

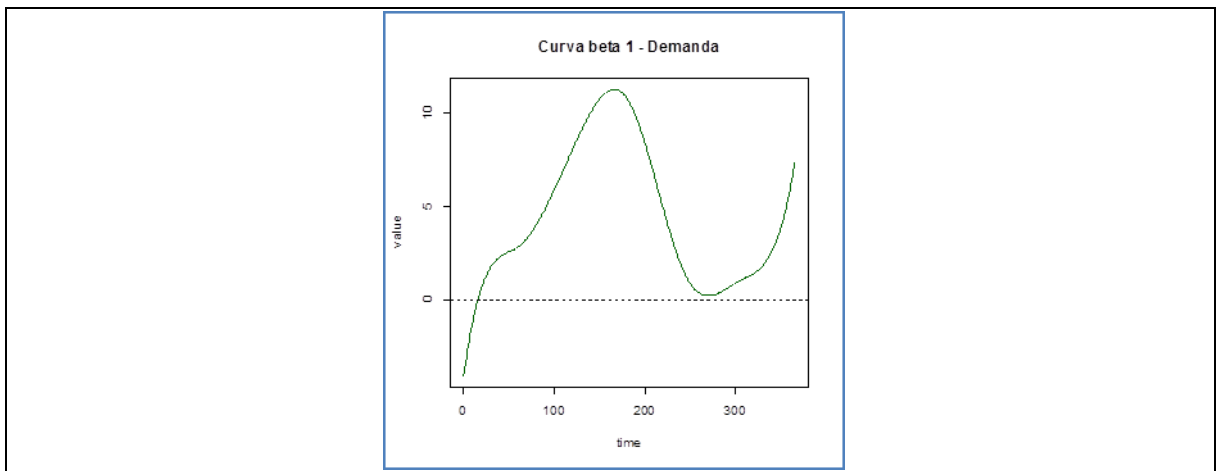
Gráfica 19. Curva beta 0.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

En este contexto no tiene interpretación el Beta 0, gráfica 19, dado que nunca la generación, la demanda y los aportes serán cero (0), esto implicaría que en un día no se transara energía eléctrica en el país.

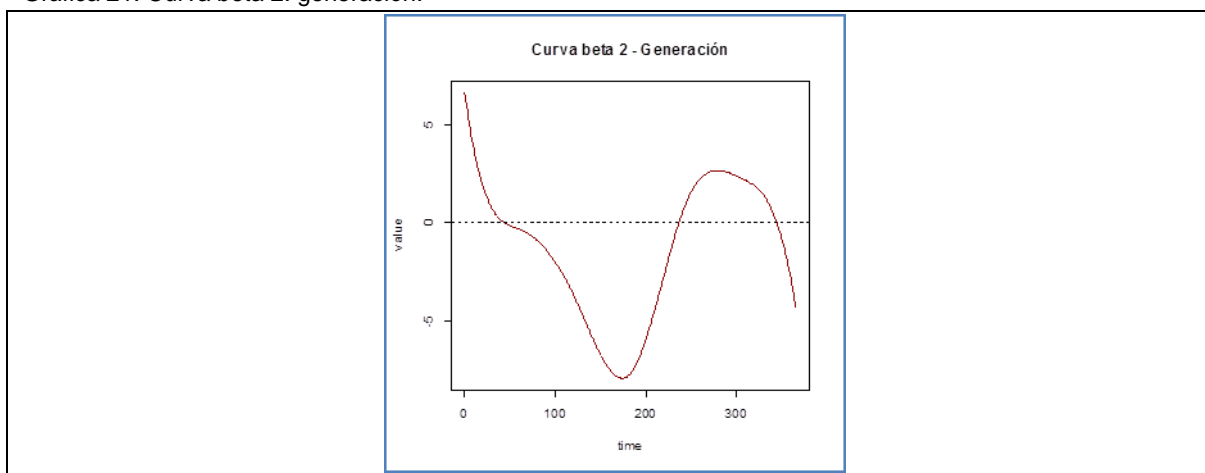
Gráfica 20. Curva beta 1: demanda.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

En cuanto a la curva Beta 1, gráfica 20, ceteris paribus, por cada GWh que aumente la demanda se estima que el precio aumente por encima de 10 \$/kWh los últimos diez días de junio en donde mayor efecto positivo presenta, seguido de los últimos quince días del año en el que se estima que por cada GWh adicional de demanda el precio tenga un aumento cercano a 7 \$/kWh mientras que ese mismo GWh adicional en demanda se estima que tenga efecto cercano a cero los últimos días de septiembre y la primera semana de octubre; así mismo se estima que tenga efecto negativo los primeros veinte días del año aproximadamente.

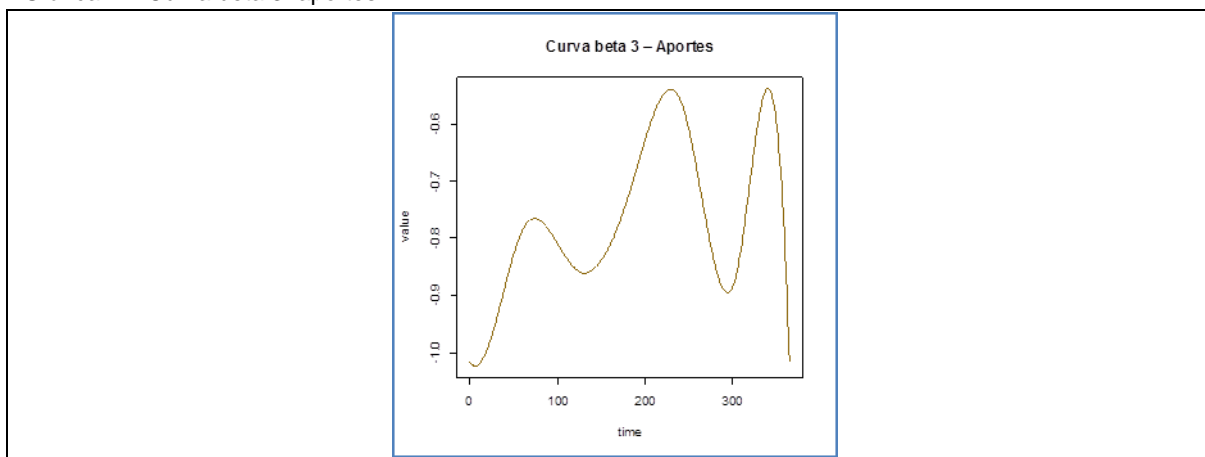
Gráfica 21. Curva beta 2: generación.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

En la curva Beta 2, gráfica 21, ceteris paribus, por cada GWh que aumente la generación se estima que el precio aumente por encima de 2 \$/kWh los primeros veinte días del año en donde mayor efecto positivo presenta, seguido del periodo entre la segunda semana de septiembre y la segunda semana de diciembre en el que se estima que por cada GWh adicional de generación el precio tenga un aumento hasta de 2 \$/kWh mientras que ese mismo GWh adicional en demanda se estima que tenga efecto negativo entre la tercera de semana de febrero y la primera semana de septiembre, siendo más negativo durante el mes de junio y las tres primeras semanas de julio aproximadamente con una caída estimada en el precio superior a -5 \$/kWh.

Gráfica 22. Curva beta 3: aportes.

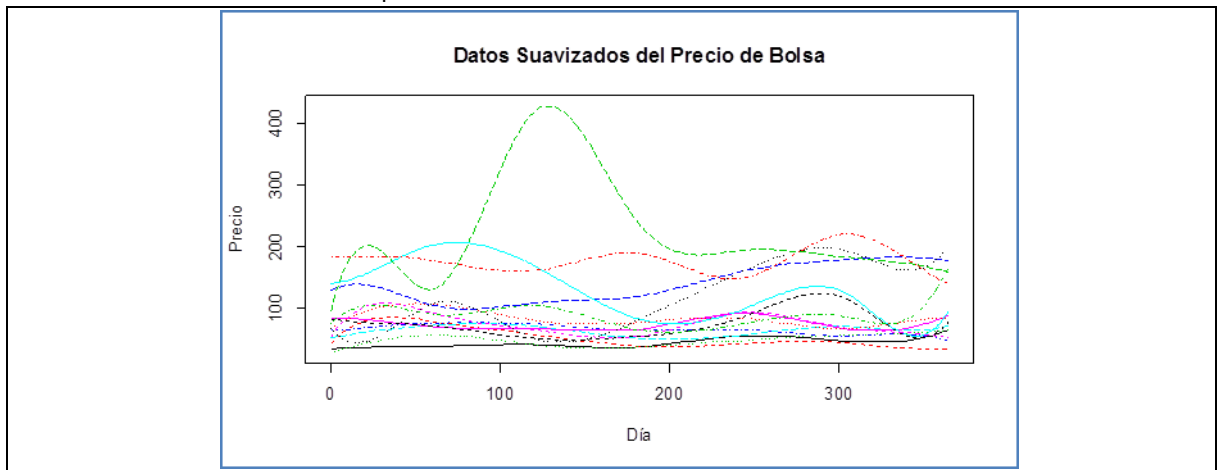


Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.



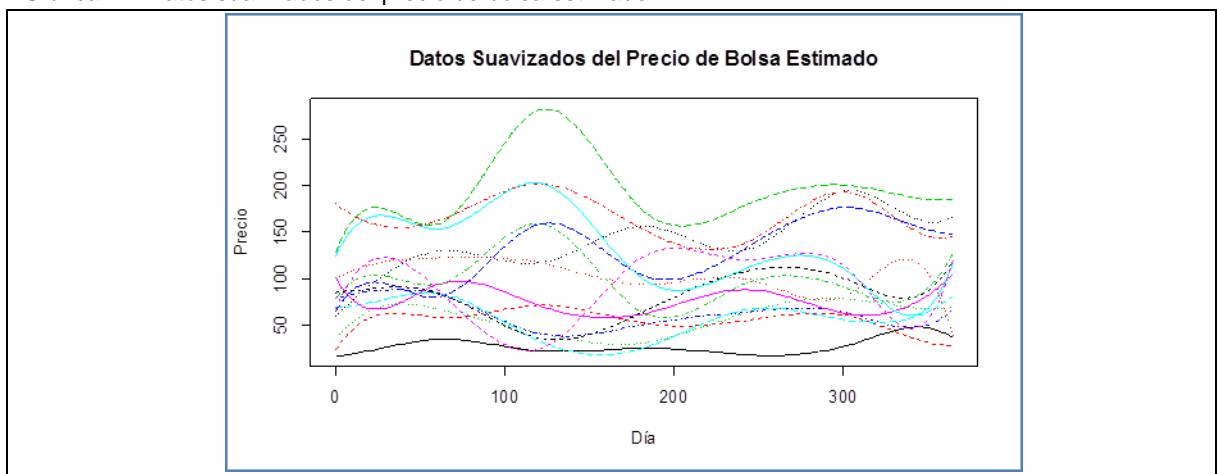
En la curva Beta 3, gráfica 22, ceteris paribus, podemos ver que por cada GWh que aumenten los aportes se estima que el precio tenga efecto negativo en un rango entre -0,6 \$/kWh y -1 \$/kWh, siendo cercano este efecto negativo a -1 \$/kWh los primeros y últimos días del año.

Gráfica 23. Datos suavizados del precio de bolsa.



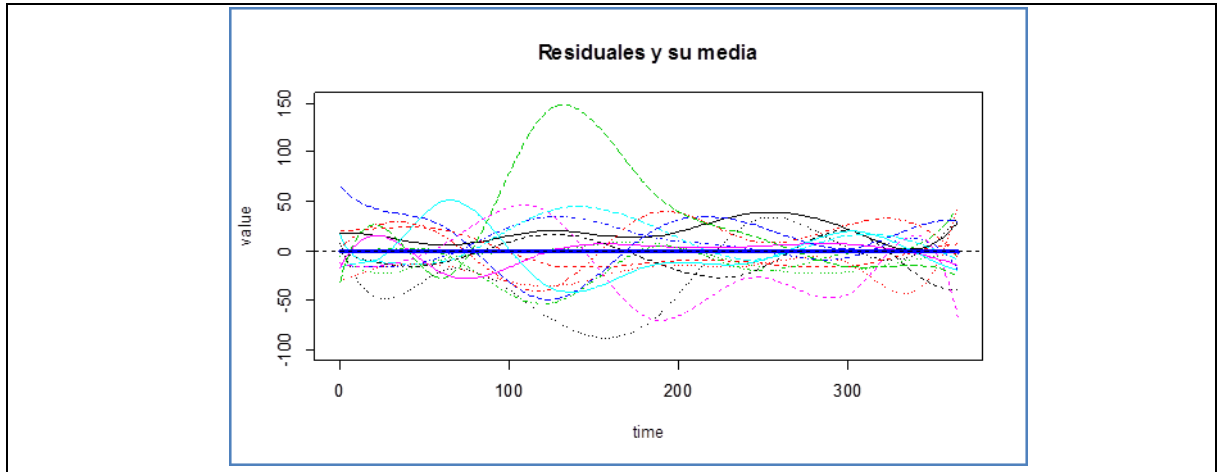
Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

Gráfica 24. Datos suavizados del precio de bolsa estimado.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

Gráfica 25. Residuales y su media.



Fuente: R CORE TEAM (2015). Gráficos obtenidos de las subrutinas ejecutadas en el programa R.

Cuando se propone un modelo para el ajuste de los datos se establecen básicamente los supuestos sobre el error de independencia, normalidad, media cero y varianza constante. Como los errores ( $\varepsilon_i$ ) son no observables, no se conocen, se pueden estimar mediante los residuales ( $\hat{\varepsilon}_i = Y_i - \hat{y}_i$ ), y hacer las pruebas de los supuestos sobre estos últimos.

Las gráficas 23 datos suavizados del precio de bolsa, 24 datos suavizados del precio de bolsa estimado y 25 residuales con su media (recta azul de mayor grosor, en el gráfico) permiten validar gráficamente que la media de los residuos es cero.

## 9. CONCLUSIONES

Algunas conclusiones que se pueden deducir de este trabajo son las siguientes:

- Tras la recolección de los datos, las variables precio de bolsa nacional de energía, generación, aportes y demanda de energía del sistema interconectado nacional se percibieron con forma de función, por ser la evolución de una variable a lo largo del tiempo (proceso estocástico), por lo tanto era posible utilizar la técnica estadística de análisis de datos funcionales, permitiendo conocer el comportamiento de las variables observadas en un mismo momento del tiempo ( $t_j$ ) en los quince años comparados.
- Con el presente estudio se pudo observar la tendencia de los últimos quince años de dichas variables a través de la observación de las mismas como datos funcionales, obteniendo la desviación estándar y la media funcionales, medidas que permitieron entender y analizar cada variable a través de su comportamiento a lo largo del año. Entre otros se pudo observar que:
  - A pesar del comportamiento alto del precio promedio de bolsa en el 2014, este no tiene un impacto muy considerable en la media funcional que oscila entre 80.00 y 120.00 \$/kWh aproximadamente, en los 365 días del año de los quince años comparados. En cuanto la desviación estándar, los precios varían más los primeros días de mayo y su menor variación se presenta los primeros días del año.
  - La variable generación tiene un comportamiento muy similar al de la variable demanda, es así como, la media funcional tiene esa misma tendencia siendo menor en los últimos y primeros días del año para ambas variables. De la misma manera, en la desviación estándar las épocas de menor variación se dan los primeros y últimos días del año tanto para la generación como para la demanda.
  - La variable aportes, presenta datos bastante dispersos con un rango amplio lo que es de entender teniendo en cuenta que dicha variable mide el aporte en GWh de las fuentes hídricas del sistema; esto se refleja en su media funcional que tiene esa misma tendencia moviéndose en un rango aproximado entre 60 y 190 GWh en promedio y cuyos picos y valles se explican por las épocas de lluvias y la

temporada seca en Colombia aunque esto puede variar por la presencia de los fenómenos de El Niño y La Niña. Para el caso de los aportes la desviación estándar nos muestra que varían más los primeros días de diciembre. Las épocas de menor variación se dan los últimos días del año y los últimos días de enero.

- A través de gráficos Bagplot se indican los periodos atípicos de las series funcionales en los periodos analizados (2000-2014) y se identifica en ellos las razones de su comportamiento a través del análisis del mercado, encontrándose que:
  - En el precio de bolsa de los 15 años observados, es el año 2014 el que presenta un comportamiento atípico y que de acuerdo al informe de XM operador del sistema, los altos precios de bolsa observados en ese año se debieron principalmente a que se presentó poca disponibilidad del parque de generación hidráulico y a la gran oleada de calor que se presentó en los primeros meses de dicho año.
  - No se identifica ningún año de los 15 años observados del periodo 2000-2014 que presente un comportamiento atípico para ni para la variable demanda, ni para la variable generación, variables que como ya se mencionó tienen comportamientos muy similares dada la naturaleza del mercado.
  - A diferencia de las anteriores, para la variable aportes, los años 2000, 2001, 2002, 2009 y 2011 presentan comportamientos atípicos, explicados por la temporada seca y de lluvias en el país y de los efectos del fenómeno de El Niño o de La Niña con sus condiciones típicas para cada año ... véase el numeral 8.4 ... identificación de series atípicas.
- A través de un modelo de regresión funcional concurrente se determina si las variables generación, aportes y demanda de energía tienen efecto en la determinación del precio de bolsa nacional de energía, hallándose que:
  - En este contexto como en muchos de la práctica el Beta 0 no tiene interpretación dado que la generación, la demanda y los aportes nunca serán cero (0) pues esto implicaría que en un día no se transara energía eléctrica en el país.

- De acuerdo a la curva Beta 1, ceteris paribus, por cada GWh que aumente la demanda se estima que el precio aumente de cero (0) a 10 \$/kWh durante casi todo el año así mismo se estima que tenga efecto negativo los primeros veinte días del año aproximadamente.
- En cuanto a la curva Beta 2, ceteris paribus, tiene el efecto contrario dónde por cada GWh que aumente la generación se estima que el precio se contraiga durante casi todo el año siendo más negativo durante el mes de junio y las tres primeras semanas de julio aproximadamente con una caída estimada en el precio superior a 5 \$/kWh, así mismo se estima que tenga efecto positivo los primeros veinte días del año aproximadamente
- Así mismo, la curva Beta 3 muestra que, ceteris paribus, por cada GWh que aumenten los aportes se estima que el precio tenga efecto negativo en un rango entre 0,6 \$/kWh y 1 \$/kWh, siendo cercano este efecto negativo a 1 \$/kWh los primeros y últimos días del año.

## 10. DISCUSIÓN

- La posibilidad de utilizar la técnica estadística de análisis de datos funcionales dio un enfoque diferente a este estudio que el posible con los datos discretos. Si bien ambos son válidos, esta técnica permitió al investigador analizar el comportamiento de la variable observada en un mismo momento del tiempo ( $t_j$ ) en los quince años comparados; por ejemplo mientras con esta técnica se obtiene que la media funcional del precio promedio de bolsa oscila aproximadamente entre 80.00 y 120.00 \$/kWh, en los 365 días del año de los quince años comparados, a través del análisis de los datos discretos como se obtuvieron originalmente, el resultado es que el precio promedio de bolsa de los quince años observados es de 97.20 \$/kWh (véase el anexo A) lo que evidencia que son enfoques diferentes en el análisis de dicha variable.
- Con el conocimiento previo del investigador, se esperaba que el alto precio promedio de bolsa presentado en el primer semestre del 2014 que superó los 400 \$/kWh ejerciera mayor influencia sobre la media funcional, sin embargo luego de obtener dicha medida se observa que oscila aproximadamente entre 80.00 y 120.00 \$/kWh, en los 365 días del año de los quince años observados, lo que refuta lo que intuitivamente se pensaba.
- Conociendo que en el mercado de energía, la oferta y la demanda deben coincidir en el tiempo debido a que la energía es un bien no almacenable se pudo comprobar que las variables generación y demanda tiene un comportamiento similar reflejado en las medidas de tendencia y dispersión analizadas.
- Así mismo, en la variable aportes conociendo que la presencia de los fenómenos de El Niño y La Niña tienen fuerte influencia en la variación de las épocas de lluvias y la temporada seca en Colombia, se podía esperar un comportamiento errático de dicha variable lo que se pudo validar en su media funcional que tiene en un rango amplio aproximado entre 60 y 190 GWh en promedio en los 365 días del año de los quince años observados.
- Se esperaba que al determinar si las variables generación, aportes y demanda de energía del sistema interconectado nacional tuvieran efecto en la determinación del precio de bolsa nacional de energía, luego de realizada

la regresión y obtenidos los betas asociados a estas variables se encuentra que:

- La curva Beta 1, ceteris paribus, nos permite validar la relación positiva que tiene la demanda en el precio dónde por cada GWh que aumente la demanda se estima que el precio aumente durante casi todo el año lo que se esperaba teniendo en cuenta que en términos generales para cualquier bien o servicio a mayor demanda mayor precio.
- La curva Beta 2, ceteris paribus, tiene el efecto contrario y también valida la relación negativa que tiene la generación en el precio dónde por cada GWh que aumente la generación se estima que el precio se contraiga durante casi todo el año lo que se esperaba teniendo en cuenta que en términos generales para cualquier bien o servicio a mayor oferta menor precio.
- La curva Beta 3, ceteris paribus, tiene el mismo efecto de la generación y también valida la relación negativa que tiene los aportes hídricos en el precio dónde por cada GWh que aumenten dichos aportes se estima que el precio se contraiga en cualquier momento del año pero en menor magnitud que la generación, lo que también se podía deducir teniendo en cuenta que a mayores aportes hídricos mayor generación hidráulica que es la generación menos costosa y de nuevo mayor oferta del servicio lo que redundaría en un menor precio.

## REFERENCIAS

ARENAS HOYOS, David. Mercado spot de energía y modelo alternativo para la fijación de un precio eficiente. Proyecto fin de carrera Profesional en Economía. Universidad EAFIT. Departamento de Economía, 2014. p. 4

COLOMBIA. COMISIÓN DE REGULACIÓN DE ENERGÍA Y GAS - CREG. Resoluciones expedidas por la CREG para regular el Mercado de Energía Eléctrico Colombiano. 1994 – 2014. Disponible en<<http://www.creg.gov.co>>

FEBRERO-BANDE, Manuel, & OVIEDO DE LA FUENTE, Manuel. Statistical Computing in Functional Data Analysis: The R Package fda.usc. Journal of Statistical Software. 2014, R package version 1.2.1 vol.51 no. 4 p. 1-28. [En línea] Disponible en<<http://www.jstatsoft.org/v51/i04/>>.

EMGESA S.A. E.S.P. Documento conceptual Generación. 2013

HYNDMAN, Rob J., & SHANG, Han Lin. Rainbow Plots, Bagplots, and Boxplots for Functional Data. Journal of Computational and Graphical Statistics. 2010, 19:1, 29-45. [En línea] Disponible en<<http://dx.doi.org/10.1198/jcgs.2009.08158>>.

LUCIO, Alejandro. ¿Es posible reducir los costos de energía eléctrica con derivados financieros? Revista Dinero [En línea] (30 de agosto de 2011 7:40:00 AM) Disponible en <<http://www.dinero.com/inversionistas/articulo/es-posible-reducir-costos-energia-electrica-derivados-financieros/134196>>

R CORE TEAM 2015. Version 3.2.2. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [En línea] Disponible en<<https://www.R-project.org/>>

RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. Second Edition United States of America: Springer Science+Business Media. Inc. 2005, p. 22-25, 43-52, 217-221, 247-260.

RAMSAY, J. O., & SILVERMAN, B. W. Functional Data Analysis. 1997 En: NAVARRO PÉREZ, Valentín. Análisis de datos funcionales: implementación y aplicaciones. Proyecto fin de carrera Licenciatura en Ciencias y Técnicas Estadísticas. Universidad Politécnica de Catalunya. Facultad de Matemáticas y Estadística. Julio 2004. p. 7



RAMSAY, J. O., WICKHAM, Hadley, GRAVES, Spencer and HOOKER, Giles. fda: Functional Data Analysis. 2014. R package version 2.4.4. [En línea] Disponible en<<http://CRAN.R-project.org/package=fda>>

SHANG, Han Lin & HYNDMAN, Rob J. rainbow: Rainbow Plots, Bagplots and Boxplots for Functional Data. 2015. R package version 3.3. [En línea] Disponible en<<http://CRAN.R-project.org/package=rainbow>>

XM S.A. E.S.P. LOS EXPERTOS EN MERCADOS. Glosario. 2014 [En línea] Disponible en<<http://www.xm.com.co>>

\_\_\_\_\_. Informe de Operación del Sistema y Administración del Mercado Eléctrico Colombiano. 2000-2014. [En línea] Disponible en<<http://www.xm.com.co>>

\_\_\_\_\_. Portal BI-Información Inteligente. Aportes Diario. 2000-2014 [En línea] Disponible en<<http://www.xm.com.co>>

\_\_\_\_\_. Portal BI-Información Inteligente. Capacidad Efectiva Neta. 2000-2014 [En línea] Disponible en<<http://www.xm.com.co>>

\_\_\_\_\_. Portal BI-Información Inteligente. Demanda Energía SIN. 2000-2014 [En línea] Disponible en<<http://www.xm.com.co>>

\_\_\_\_\_. Portal BI-Información Inteligente. Generación. 2000-2014 [En línea] Disponible en<<http://www.xm.com.co>>

\_\_\_\_\_. Portal BI-Información Inteligente. Precio Bolsa Nacional. 2000-2014 [En línea] Disponible en<<http://www.xm.com.co>>

\_\_\_\_\_. Términos legales uso Portal BI XM. 2014 [En línea] Disponible en<<http://www.xm.com.co>>

## **ANEXOS**

## ANEXO A. ESTADÍSTICOS PRINCIPALES EN LOS DATOS DISCRETOS RESULTADOS EN R

```
##Estadísticos generales
```

```
datos=read.csv2("BD.csv",header=T, sep=";",dec=".")
```

```
attach(datos)
```

```
summary (datos)
```

Precio	Demanda	Generación	Aportes
Min. : 28.84	Min. : 86.79	Min. : 86.53	Min. : 25.25
1st Qu.: 57.26	1st Qu.:126.75	1st Qu.:127.83	1st Qu.: 86.92
Median : 76.21	Median :141.80	Median :145.32	Median :124.33
Mean : 97.20	Mean :142.44	Mean :144.72	Mean :133.98
3rd Qu.:120.52	3rd Qu.:158.12	3rd Qu.:160.96	3rd Qu.:168.13
Max. :478.87	Max. :255.39	Max. :193.36	Max. :591.25